دانشگاه صنعتی شریف

دانشکده مهندسی برق

پایان نامه دکتری

گرایش مخابرات سیستم

نمونه برداری فشرده با روش‌های غیرتصادفی

نگارنده

آرش امینی

استاد راهنما

دکتر فرخ مروستی

بهمن ۱۳۸۹

این پروژه تحت قرارداد پژوهشی شماره ۵۰۰/۶۹۴۷ مورخ ۱۳۸۸/۲/۱۴ از پشتیبانی مادی و معنوی

مرکز تحقیقات مخابرات ایران بهره‌مند شده است.

بسم الله الرحمن الرحيم

توجه

این پروژه بر اساس قرارداد شماره (۵۰۰/٦۹٤۷) از حمایت مالی مرکز تحقیقات مخابرات ایران برخوردار شده است.

بسمه تعالی

دانشگاه صنعتی شریف
دانشکده مهندسی برق

پایان‌نامه دکتری

عنوان: **نمونه‌برداری فشرده با روش‌های غیرتصادفی**

نگارش: **آرش امینی**

اعضا هیات داوران:

دکتر فرخ مروستی امضاء:.....

دکتر محمدرضا عارف امضاء:.....

دکتر جواد صالحی امضاء:.....

دکتر پاییز عزمی امضاء:.....

دکتر بابک سیف امضاء:.....

دکتر کسری علیشاهی امضاء:.....

Dr. Michael Unser امضاء:.....

تاریخ: ۲۰ بهمن ۱۳۸۹.

تقدیم و قدردانی

طی مراحل اجرای این پایان‌نامه از مساعدت و همراهی بسیاری از دوستان بهره‌مند بودم که بر خود لازم می‌دانم در این‌جا از آن‌ها تشکر کنم. ترتیب ذکر اسامی این افراد، الفبایی است:

علی اخایی، امیر اکبری، پدرام پاد، پوریا پاکروح، امین توکلی، آرین حاذقی، فرزان حدادی، احمدرضا حسینی، پویا دهقان‌تفتی، سینا زاهدپور، علی سالمی، مهدی سلطان‌الکتابی، مجتبی سلطان‌علیان، میلاد شریف، مهدی عالم، محسن غفاری، محمود فردوسی‌زاده، فرشید فرهت، سهیل فیض‌خانکندی، امین کرباسی، رامتین مدنی، وحید منتظرحجت، پیمان مهاجرین‌اصفهانی و مجید ولی‌الله‌زاده.

نهایت سپاس خود را نثار خانواده‌ام می‌کنم که همواره پشتیبان و مشوقم بوده‌اند؛ به‌ویژه مادرم، مهران مجدی، که زحمت تایپ و ویراستاری این پایان‌نامه را بر عهده داشت.

از جناب آقای دکتر بابک سیف به دلیل مطالعه دقیق و بهبود کیفیت متن پایان‌نامه، قدردانی می‌کنم.

موفقیت خود را در اتمام دوره دکتری مدیون راهنمایی‌ها و حمایت‌های جناب آقای دکتر فرخ مروستی هستم که مرا همچون فرزند خویش مورد لطف و محبت قرارداد.

و در انتها، این پایان‌نامه را به همسر عزیز و فداکارم، میترا فاطمی، تقدیم می‌کنم که در طول دوره دکتری، نه تنها مرا همراهی نمود، بلکه در بسیاری از موارد با از خودگذشتگی مرا مورد حمایت خویش قرارداد.

چکیده:

در مبحث نمونه برداری فشرده، به دنبال نمایش سیگنال‌های تنک با کمترین تعداد نمونه ممکن هستیم به نحوی که نمونه‌ها به طور یکتا بردار تنک را مشخص کنند. در این پایان‌نامه، مسأله نمونه‌برداری فشرده با ساختار معین (در مقابل تصادفی) از چندین دیدگاه مورد بررسی قرار می‌گیرد. نحوه نمونه‌برداری در تمام روشهای ارائه شده تا کنون به صورت خطی و توسط یک ماتریس که عمدتاً تصادفی است، در نظر گرفته شده است. در این پایان‌نامه ابتدا چند روش تولید ماتریس معرفی می‌شود که یکی از آن‌ها یک روش مبتنی بر کدهای متعامد نوری برای ساخت ماتریس‌های باینری است. در روشی دیگر، به کمک کدهای BCH راه‌کاری برای تولید ماتریس‌های دو قطبی ارائه می‌دهیم. کدهای بلوکی گردشی علاوه بر آن که قابلیت ساخت ماتریس نمونه‌برداری را فراهم می‌کنند، پیچیدگی محاسباتی در روند بازسازی سیگنال تنک را نیز کاهش می‌دهند. همچنین به کمک تعمیم کدهای BCH، ماتریس‌های مختلط را معرفی می‌کنیم. این ماتریس‌ها از نظر ابعاد تنوع بیشتری ایجاد می‌کنند و در نتیجه راحت‌تر می‌توان آن‌ها را با ساختارهای واقعی منطبق کرد. علاوه بر طرح‌های تولید ماتریس، در گام بعدی روش‌هایی برای ادغام این ماتریس‌ها ارائه می‌کنیم به گونه‌ای که ماتریس حاصل همچنان شرایط لازم ماتریس‌های نمونه‌برداری را داراست و به علاوه طیف وسیع‌تری از ابعاد را پوشش می‌دهد.

جدا از مبحث نمونه‌برداری خطی، روش های غیرخطی و مزایای آن‌ها را نیز بررسی می‌کنیم. در صورتی که حوزه تنک‌بودن برای سیگنال‌های مورد نظر مشخص نباشد، هیچ روش خطی نمی‌تواند بازسازی کامل را تضمین کند. در این جا نشان می‌دهیم که در چنین شرایطی می‌توان به کمک روش‌های غیرخطی سیگنالهای تک-تنک را بازسازی کرد. علاوه بر این در صورت مشخص بودن حوزه تنک‌بودن، می‌توان به کمک روش‌های غیرخطی، همزمان تعداد نمونه‌های لازم و پیچیدگی محاسباتی در نحوه بازسازی را به طور قابل ملاحظه‌ای کاهش داد. علی‌رغم تمام مزایای روش‌های غیرخطی، مشکل اصلی این روش‌ها حساسیت نسبت به نویز جمعی است.

در مبانی نمونه‌برداری فشرده، هدف ادغام دو عمل نمونه‌برداری و فشرده‌سازی است. از این رو، فرض واقعی‌تر آن است که سیگنال‌های ورودی پیوسته‌اند و نه گسسته. اما اولین مشکل در این راه، تعریف مفاهیم تنک‌بودن و فشرده‌پذیری برای سیگنال‌های پیوسته است. در این پایان‌نامه نشان می‌دهیم که این دو مفهوم را می‌توان به طور مناسبی به دنباله‌های نامتناهی یقینی و حتی تصادفی تعمیم داد. این مرحله در حقیقت اولین گام برای تعمیم مباحث نمونه‌برداری فشرده به مجموعه‌های بی‌نهایت بعدی و پیوسته از سیگنال‌هاست.

همان طور که اشاره شد، مفهوم تنک‌بودن اغلب بر روی بردارها مورد بررسی قرار می‌گیرد. اما پیشتر تعمیم مفهوم تنک‌بودن برای ماتریس بر اساس رتبه آن (به جای مقادیر درایه‌ها) نیز مورد توجه قرار گرفته است. در انتهای پایان‌نامه به بررسی مسأله خاصی در مورد ماتریس‌های تنک (به مفهوم رتبه) می‌پردازیم. به طور دقیق‌تر، بررسی می‌کنیم که چگونه می‌توان یک ماتریس کم‌رتبه را که درایه‌های آن تحت تاثیر یک اعوجاج غیرخطی یکسان قرار گرفته‌اند، بازیابی و در نتیجه فشرده کرد.

کلمات کلیدی:

۱- نمونه‌برداری فشرده Compressed Sensing.
۲- تنک‌بودن Sparsity.
۳- تصویر خطی Linear Projection.
۴- نمونه‌برداری غیرخطی Nonlinear Sampling.
۵- دنباله تصادفی با توزیع مستقل یکنواخت i.i.d. Sequence.
۶- فشرده‌پذیری Compressibility.
۷- ماتریس کم‌رتبه Low-rank Matrix.

فهرست مطالب

۱ مقدمه **۱**

۱–۱ نمونه‌برداری شانون تا نمونه‌برداری فشرده ۱

۱–۲ نمونه‌برداری فشرده ۳

۱–۳ کاربردهای نمونه‌برداری فشرده ۸

۱–۴ تکمیل ماتریس ۹

۱–۵ طبقه بندی مطالب پایان‌نامه ۱۰

۲ کلیات نمونه‌برداری فشرده **۱۲**

۲–۱ مقدمه ۱۲

۲–۲ عرض گلفاند ۱۲

۲–۳ بررسی نمونه‌برداری فشرده از دیدگاه نظریه نرخ–اعوجاج ۱۶

۲–۴ ارتباط نمونه‌برداری فشرده با نظریه کدگذاری ۱۸

۲–۵ مروری بر روش‌های بازسازی ۲۱

۲–۵–۱ کمینه کردن نرم ℓ_1 . . . . . ۲۱

۲–۵–۲ روش‌های حریص ۲۳

۲–۵–۳ روش‌های آستانه‌ای ۲۴

۲–۵–۴ روش‌های تقریب نرم ۲۵

۳ مروری بر ماتریس‌های حسگر ۲۷

۳-۱ مقدمه ۲۷

۳-۲ ماتریس‌های حسگر تصادفی ۲۸

۳-۳ ماتریس‌های حسگر غیرتصادفی ۳۱

۴ ماتریس‌های حسگر غیرتصادفی پیشنهادی ۳۹

۴-۱ مقدمه ۳۹

۴-۲ ماتریس‌های دودویی ۴۰

۴-۳ ماتریس‌های دو قطبی به کمک کدهای BCH ۴۳
۴-۳-۱ کدهای BCHبا $\tilde{d}_{min}$ بزرگ ۴۴
۴-۳-۲ الگوریتم تولید ماتریس ۴۸

۴-۴ ماتریس‌های مختلط به کمک کدهای غیردودویی ۵۰
۴-۴-۱ کدهای pBCH –سمبلی با $\tilde{d}_{min}$ بزرگ ۵۳
۴-۴-۲ الگوریتم تولید ماتریس ۵۶

۴-۵ ادغام ماتریس‌ها ۵۷
۴-۵-۱ ادغام با ماتریس‌های دودویی ۵۸
۴-۵-۲ ضرب کرونکر ۶۰

۴-۶ بازسازی سریع ۶۱

۴-۷ نتایج شبیه‌سازی ۶۴

۵ نمونه‌برداری غیرتصادفی غیرخطی ۷۰

۵-۱ مقدمه ۷۰

۵-۲ نمونه‌برداری غیرتصادفی برای حوزه تنک‌بودن نامعلوم ۷۰

۵-۳ نمونه‌برداری غیرتصادفی برای حوزه تنک‌بودن معلوم ۷۶

۵-۴ نتایج شبیه‌سازی ۸۱

۶ دنباله‌های تصادفی فشرده‌پذیر ۸۳

۶-۱ مقدمه ۸۳

۶-۲ دنباله‌های نامتناهی فشرده‌پذیر ۸۴

۶-۳ فشرده‌پذیری دنباله‌های تصادفی ۸۶

۶-۴ تشخیص توزیع احتمال‌های فشرده‌پذیر ۸۸

۶-۵ نتایج عددی ۹۶

۷ کاهش رتبه در ماتریس‌ها به کمک عملگرهای المانی ۹۹

۷-۱ مقدمه ۹۹

۷-۲ نمادها و شرح مساله ۱۰۲

۷-۳ توان صحیح ۱۰۳

۷-۴ توان حقیقی ۱۰۵

۷-۵ نتایج عددی ۱۰۹

۸ جمع بندی و نتیجه‌گیری ۱۱۲

الف محاسبه $\tilde{k}$ ۱۱۵

ب محاسبه δ ۱۱۸

فهرست جداول

۱-۴ چندجمله‌ای آزمون توازن برای مقادیر متفاوت $\tilde{m}$ و $i = ۳$. ۵۰

۲-۴ $\frac{\mu_{BCH}}{\mu_{WB}}$ برای چند ماتریس p–سمبلی $p^{۳l} \times (p^{۲l} - ۱)$ با pهای مختلف. ۵۶

۱-۶ تعدادی از توزیع‌های فشرده‌پذیر با میرایی مرتبه $|t|^{-(q+۱)}$ ۹۵

فهرست اشکال

۱–۱ (a) یک بردار تنک در فضای سه بعدی، (b) تقریب حاصل از کمینه‌سازی نرم ℓ_2 برای این بردار، (c) تقریب حاصل از کمینه‌سازی نرم ℓ_1 برای این بردار [۱۰]. ۶

۱–۴ درجه چندجمله‌ای $h(x)$ برای مقادیر متفاوت $\tilde{m}$ و i. ۵۰

۲–۴ عمل ادغام ماتریس $\mathbf{B}$ با ماتریس دودویی $\mathbf{A}$. ۵۹

۳–۴ ضرب کرونکر دو ماتریس ($\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$). ۶۰

۴–۴ درصد بازسازی ($\text{SNR}_{rec.} \geq ۱۰۰dB$) در مقادیر مختلف k. ماتریس‌های حسگر با ساختار OOCو Devoreابعاد ۳۷۸ × ۶۳ و ۳۷۸ × ۶۴ دارند. همچنین دو ماتریس تصادفی دودویی با ابعاد ۳۷۸ × ۶۴ و ۳۷۸ × ۶۳ در نظر گرفته شده‌اند. ۶۴

۵–۴ درصد بازسازی ($\text{SNR}_{rec.} \geq ۱۰۰dB$) در مقادیر مختلف k. ماتریس‌های حسگر با ساختار BCH، Devoreو ادغامی (Ternary) ابعاد ۵۱۲ × ۶۳، ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ دارند. همچنین دو ماتریس تصادفی گوسی ابعاد ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ دارند. ۶۵

۶–۴ SNRسیگنال‌های ۱۵–تنک بازسازی شده هنگامی که نمونه‌های فشرده تحت تاثیر نویز جمعی با توان‌های متفاوتی قرار گیرند. ماتریس‌های حسگر با ساختار BCH، Devoreو ادغامی (Ternary) ابعاد ۵۱۲ × ۶۳، ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ دارند. همچنین دو ماتریس تصادفی گوسی ابعاد ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ دارند. ۶۶

۷–۴ درصد بازسازی کامل ($\text{SNR}_{rec.} \geq ۱۰۰dB$) هنگامی که نمونه‌ها بدون نویز هستند. ماتریس BCHبر مبنای کدهای $p = ۳$ سمبلی ساخته شده است و ابعاد تمامی ماتریس‌ها ۷۲۹ × ۸۰ است. ۶۷

۸–۴ SNRسیگنال ۲۵–تنک بازسازی شده برای توان‌های متفاوت نویز جمعی. ماتریس BCHبر مبنای کدهای $p = ۳$ سمبلی ساخته شده است و ابعاد تمامی ماتریس‌ها ۷۲۹ × ۸۰ است. . . ۶۸

۹–۴ SNRسیگنال ۲۵–تنک بازسازی شده برای توان‌های متفاوت نویز جمعی. ماتریس BCHبر مبنای کدهای $p = ۵$ سمبلی ساخته شده است و ابعاد تمامی ماتریس‌ها ۱۵۶۲۵ × ۶۲۴ است. ۶۸

۱۰–۴ درصد بازسازی کامل ($\text{SNR}_{rec.} \geq ۱۰۰dB$) هنگامی که نمونه‌ها بدون نویز هستند. ابعاد ماتریس‌های بر مبنای ادغام دودویی، ادغام کرونکر، توابع Chirp، تصادفی گوسی و سطرهای تصادفی ماتریس DFTبه ترتیب عبارتند از ۴۶۰۸ × ۶۴، ۱۷۲۸ × ۶۳، ۴۶۰۸ × ۷۵، ۴۶۰۸ × ۶۴ و ۴۶۰۸ × ۶۴. . . . . ۶۹

۴-۱۱ مقایسه زمان لازم برای بازسازی یک بردار تنک ۱×۱۵۶۲۵ از نمونه‌های فشرده ۱×۶۲۴ توسط ماتریس تصادفی (OMPساده) و ماتریس‌های BCH$۵ = p$ سمبلی (OMPتسریع شده). ۶۹

۵-۱ کیفیت سیگنال بازسازی شده توسط نمونه‌های غیرخطی برای مرتبه‌های متفاوت تنک‌بودن. . ۸۱
۵-۲ کیفیت سیگنال بازسازی شده توسط نمونه‌های غیرخطی به ازای $۶ = k$ در OSRهای متفاوت. ۸۲

۶-۱ یک تحقق زمانی از دنباله تصادفی و .i.i.dبا توزیع احتمال Student's tو پارامتر $۰/۵ = q$. . ۹۶
۶-۲ نمودارهای متوسط تابع $\frac{\sigma_p(k;\{a_i\}_{i=۱}^n)}{\sigma_p(n;\{a_i\}_{i=۱}^n)}$ بر حسب $\frac{k}{n}$ برای تحقق‌های .i.i.dاز توزیع احتمال‌های گوسی، لاپلاس، کوشی و Student's t($q = ۰/۵$) به ازای $۱ = p$ و $n = ۱۰^۴$. . ۹۷
۶-۳ نمودارهای متوسط تابع $\frac{\sigma_p(k;\{a_i\}_{i=۱}^n)}{\sigma_p(n;\{a_i\}_{i=۱}^n)}$ بر حسب $\frac{k}{n}$ در nهای مختلف برای تحقق‌های .i.i.dاز توزیع احتمال کوشی هنگامی که $۰/۹ = p$ و $۱/۱ = p$ مورد نظر باشند. ۹۸

۷-۱ آرایش دایروی حسگرها در توموگرافی فراصوتی. ۱۰۱
۷-۲ (a)ضرایب بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ و (b)مقادیر چندجمله‌ای‌های حاصل از در نظر گرفتن متناهی جمله بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ پس از حذف فاکتور $x^۴$ هنگامی که نویز جمعی در حد $\text{SNR} = ۱۰۰dB$ است. ۱۱۰
۷-۳ (a)ضرایب بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ و (b)مقادیر چندجمله‌ای‌های حاصل از در نظر گرفتن متناهی جمله بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ پس از حذف فاکتور $x^۴$ هنگامی که نویز جمعی در حد $\text{SNR} = ۵۰dB$ است. ۱۱۰

الف-۱ مقادیر بدون تقریب $\kappa_b^{(a)}$ در چند انتخاب متفاوت از a و b. ۱۱۷

فهرست کلمات اختصاری

BP	Basis Pursuit
CFAR	Constant False Alarm Rate
CS	Compressed Sensing
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
FISTA	Fast ISTA
GPSR	Gradient Projection for Sparse Reconstruction
IDFT	Inverse Discrete Fourier Transform
IHT	Iterative Hard Thresholding
IMAT	Iterative Method with Adaptive Thresholding
ISTA	Iterative Shrinkage-Thresholding Algorithm
LS	Least Squares
MP	Matching Pursuit
OMP	Orthogonal Matching Pursuit
OOC	Optical Orthogonal Code
OSR	Over-Sampling Ratio
RIP	Restricted Isometry Property
RS	Reed-Solomon
SL0	Smoothed ℓ_0 norm
SPGL1	Spectral Projected-Gradient ℓ_1 norm
SNR	Signal to Noise Ratio
SVD	Singular Value Decomposition

فصل ۱

مقدمه

در این بخش ابتدا تاریخچه و کلیاتی در مورد مبحث نمونه‌برداری فشرده ارائه می‌شود. همچنین تعاریف کلیدی که در اکثر فصل‌ها مورد استفاده قرار می‌گیرند، در اولین بخش گنجانده شده است. در انتهای این فصل، نحوه دسته‌بندی مطالب در پایان‌نامه به طور مختصر شرح داده خواهد شد.

۱-۱ نمونه‌برداری شانون تا نمونه‌برداری فشرده

در قضایای نمونه‌برداری کلاسیک (که در ادبیات مهندسی برای اولین بار توسط شانون [۸۴] مطرح شد ولی تاریخچه‌ای قدیمی‌تر دارد [۹۸])، به دنبال بیان یک سیگنال باند محدود توسط نمونه‌های زمانی آن هستیم. به لحاظ پیشینه تاریخی، ابتدا سیگنال‌های پایین‌گذر[۱] مورد بررسی قرار گرفتند. شایان ذکر است که مفاهیم باند محدود، پایین گذر، بالاگذر و ... منوط به تعریف حوزه‌ای به نام حوزه فرکانس است که در بررسی‌های کلاسیک نمونه‌برداری، این حوزه همان تبدیل فوریه فرض شده‌است؛ به بیان بهتر، تاکنون دو مفهوم حوزه فرکانس و تبدیل فوریه به یک معنا به کار رفته‌اند که در متن پیش رو لزوماً یکسان نخواهند بود. مفهوم حوزه فرکانس در ادامه شرح داده خواهد شد. بررسی‌های انجام شده در نمونه‌برداری یکنواخت از سیگنال‌های پایین گذر مبین آن است که حداقل نرخ نمونه‌برداری برای قابلیت بازسازی سیگنال اصلی، دو برابر پهنای باند تبدیل فوریه است (نرخ نایکوئیست)[۷۵]. نتیجه مذکور به سیگنال‌های میان‌گذر [۱۸، ۳۴] و چندبانده [۱۷] و نمونه‌برداری‌های غیریکنواخت [۶۹] و حتی تصادفی [۳۲] نیز تعمیم داده شده است. در تمام موارد فوق، کمترین

[۱] Lowpass

نرخ نمونه‌برداری که بازسازی کامل سیگنال را تضمین کند، همواره ضریبی از پهنای باند است (این ضریب در بهترین حالت که نمونه‌برداری غیریکنواخت تطبیق‌شده بر سیگنال استفاده شود، برابر یک است!). در مقالات اخیر [۱۰، ۲۶، ۴۰] شرط پهنای‌باند محدود، با شرط قوی‌تری (محدود کننده‌تر) جایگزین شده و نتایج جالبی بدست آمده است. مخلوطی از k سیگنال سینوسی را فرض کنید که با استفاده از نمونه‌های زمانی آن می‌خواهیم فرکانس و ضرایب این سینوس‌ها را (در حالت کلی، ضرایب مختلط که فاز را نیز در بر گیرد) تعیین کنیم. در صورتی که بخواهیم از قضایای کلاسیک نمونه‌برداری استفاده کنیم، نمونه‌برداری یکنواخت را باید با نرخی بیشتر از دو برابر بزرگترین فرکانس موجود در بین سیگنال‌های سینوسی انجام داد؛ یعنی نرخ نمونه‌برداری بدون توجه به تعداد سیگنال‌های تک‌فرکانس و تنها براساس بزرگترین فرکانس تعیین می‌شود. در حالت حدی فرض کنیم تنها یک سیگنال سینوسی $\alpha e^{j۲\pi ft}$ موجود باشد؛ به وضوح تنها با دو نمونه زمانی می‌توان α و f را تعیین کرد، حال آن که با استفاده از نرخ نایکوئیست، بی‌نهایت نمونه با فاصله‌های زمانی $\Delta t = \frac{۱}{۲f}$ مورد نیاز است. اگر سیگنال مجموع چند سینوسی را به صورت یک سیگنال چندبانده با باندهای کم‌عرض فرض کنیم، برای بازسازی آن به نمونه‌برداری غیریکنواخت تناوبی[۲] نیازمندیم که فاصله زمانی بین نمونه‌ها براساس مکان باندهای فرکانسی (تبدیل فوریه) تعیین می‌شوند؛ از آن‌جا که فرکانس سیگنال‌های سینوسی نامعلوم است، تعیین این پارامترها هم ممکن نیست. پس روشن است که قضایای نمونه‌برداری کلاسیک در مورد سیگنال‌های مشابه مثال ذکر شده، نرخ‌های نمونه‌برداری غیربهینه‌ای پیشنهاد می‌کنند. مثال مطرح شده درخصوص مجموع چند سیگنال سینوسی، در مسائل مربوط به تخمین طیف نیز اهمیت بسیاری دارد؛ از جمله روش‌های معروف پیشنهادی در تخمین طیف برای یافتن اندازه و فرکانس سیگنال‌های سینوسی، می‌توان به روش‌های Prony [۳۶]، Pisarenko [۷۸] و MUSIC [۸۳] اشاره کرد. در تمام این روش‌ها، ابتدا ماتریس خودهمبستگی سیگنال از روی نمونه‌ها تخمین زده می‌شود و سپس برحسب مقادیر و بردارهای ویژه، پارامترهای مورد نظر تعیین می‌شوند. روشن است که برای تخمین مناسبی از ماتریس خودهمبستگی، تعداد زیادی نمونه لازم است تا میانگین‌های مربوط، به مقادیر امید ریاضی نزدیک شوند. نکته مثبت این روش‌ها، مرتبط بودن ابعاد ماتریس خودهمبستگی مورد نیاز به تعداد سیگنال‌های تک‌فرکانس است. به عبارت دیگر، با کاهش و یا افزایش تعداد مولفه‌های تک‌فرکانس، تعداد نمونه‌های مورد نیاز و پیچیدگی محاسباتی نیز به ترتیب کم و زیاد می‌شوند (برخلاف روش کلاسیک استفاده از

[۲] Periodic Nonuniform Sampling

نرخ نایکوئیست). در روش جدید نمونه‌برداری که به نمونه‌برداری فشرده[۳] معروف شده است، هدف کاهش تعداد نمونه‌های لازم برای بازسازی سیگنال‌های مشابه با مثال مطرح شده است که در حوزه‌ای به نام حوزه‌ی فرکانس (که در مثال ذکر شده همان تبدیل فوریه است) نمایش تنک[۴] داشته باشند. به بیان دیگر، باید در حوزه‌ی فرکانس مورد نظر، تعداد ضرایب غیرصفر به مراتب کمتر از تعداد ضرایب صفر باشد. گفتنی است که در مورد کنار هم قرار گرفتن ضرایب ناصفر، فرضی وجود ندارد؛ به همین دلیل پهنای‌باند در این حالت مصداق نخواهد داشت.

۱-۲ نمونه‌برداری فشرده

مبحث نمونه‌برداری فشرده که ابتدا در مقالات [۲۳، ۲۵، ۴۰] معرفی شد و هم‌اکنون جزء موضوعات روز تحقیق به شمار می‌رود، حاصل از بکارگیری شرط تنک‌بودن در مسأله‌ی نمونه‌برداری است. نکته‌ی جالب در این مبحث آن است که قبل از پایه‌ریزی روش‌های نمونه‌برداری، روش بازسازی کاملاً شناخته شده بود. در حقیقت موفقیت چشمگیر روش کمینه کردن نرم ℓ_1 موجب طراحی روش‌های نمونه‌برداری منطبق با این روش شد.

فرض کنید بردار $\mathbf{x}_{n\times 1}$ مبین یک سیگنال گسسته و متناهی در زمان باشد. گوییم $\mathbf{x}_{n\times 1}$ یک سیگنال k-تنک است اگر نمایش این بردار در یک حوزه متعامد یکه، حداکثر k مولفه ناصفر داشته باشد. به بیان ریاضی:

$$\mathbf{x}_{n\times 1} = \mathbf{\Psi}_{n\times n} \cdot \mathbf{s}_{n\times 1} \qquad (۱-۱)$$

به طوری که $\mathbf{\Psi}_{n\times n}$ یک ماتریس یکانی (معرف حوزه متعامد یکه) و $\mathbf{s}_{n\times 1}$ برداری با حداکثر k درایه ناصفر است. مثلاً اگر $\mathbf{x}_{n\times 1}$ تصویری باشد که به نحوی به بردار تبدیل شده باشد و $\mathbf{\Psi}_{n\times n}$ معادل ماتریسی باشد که عکس تبدیل DCT تصویر را ایجاد کند (به دلیل خطی بودن تبدیل، حتماً چنین ماتریسی وجود دارد)، $\mathbf{s}_{n\times 1}$ برداری است که با تقریب نسبتاً خوبی، تنک فرض می‌شود ($\mathbf{s} = \mathbf{\Psi}^{-1} \cdot \mathbf{x}$). در واقع، این حقیقت ایده اصلی چندین روش فشرده‌سازی تصاویر مثل JPEG 2000 است. حال فرض کنید بخواهیم بردار $\mathbf{x}_{n\times 1}$ را با کمترین اطلاعات (از نظر تعداد نمونه) نمایش دهیم؛ از آنجا که $\mathbf{s}_{n\times 1}$ تنها k مقدار ناصفر دارد، با در اختیار داشتن مقادیر و مکان‌های ناصفر بردار $\mathbf{s}$، این بردار به طور یکتا مشخص می‌شود و در صورت آگاهی از ماتریس $\mathbf{\Psi}$، بردار $\mathbf{x}_{n\times 1}$ نیز به طور یکتا مشخص خواهد شد. در نتیجه، نمایش بردار $\mathbf{x}_{n\times 1}$ تنها با $2k$ نمونه (k نمونه برای اندیس

[۳] Compressed Sensing

[۴] Sparse

مکان‌های ناصفر و k نمونه برای مقادیر این مکان‌ها) امکان پذیر است. اما مشکل اینجاست که برای بدست‌آوردن چنین نمایشی، باید ابتدا کل بردار $\mathbf{x}_{n\times 1}$ را در دست داشته باشیم، سپس تبدیل آن را در حوزه فرکانس بدست آوریم و مکان‌های ناصفر (و یا مکان‌های با مقادیر قابل توجه) آن را مشخص کنیم. پس، ابتدا باید سیگنال به طور کامل ذخیره شود و سپس با توجه به تنک‌بودن آن در حوزه‌ی فرکانس مورد نظر، فشرده شود.

در نمونه‌برداری فشرده، بر خلاف آن چه گفته شد، مایلیم که نمونه‌برداری به صورت غیروفقی[۵] و خطی صورت گیرد. به عبارت دیگر اگر بردار نمونه‌ها را با $\mathbf{y}_{m\times 1}$ نشان دهیم، داریم:

$$\mathbf{y}_{m\times 1} = \mathbf{\Phi}_{m\times n} \cdot \mathbf{x}_{n\times 1} = \mathbf{\Phi}_{m\times n} \cdot \mathbf{\Psi}_{n\times n} \cdot \mathbf{s}_{n\times 1} \qquad (۲-۱)$$

که ماتریس $\mathbf{\Phi}_{m\times n}$ مستقل از $\mathbf{s}_{n\times 1}$ انتخاب شده است و به ماتریس حسگر[۶] شهرت دارد. در حقیقت ابعاد ماتریس حسگر به نوعی نرخ فشرده‌سازی را مشخص می‌کنند؛ اگر نمونه‌برداری بدون از دست رفتن اطلاعات صورت گرفته باشد، یک بردار n بعدی به یک بردار m بعدی معادل تبدیل شده است. برای بازسازی بردار $\mathbf{x}$ و یا بردار تنک $\mathbf{s}$ که با $\mathbf{x}$ معادل است، با معادله فرو-معین[۷] ذیل روبرو هستیم:

$$\mathbf{y}_{m\times 1} = \mathbf{\Phi}_{m\times n} \cdot \mathbf{\Psi}_{n\times n} \cdot \mathbf{s}_{n\times 1} = \mathbf{A}_{m\times n} \cdot \mathbf{s}_{n\times 1} \qquad (۳-۱)$$

که در حالت $m < n$ بی‌شمار جواب دارد ($\mathbf{y}_{m\times 1}$ به عنوان بردار معلوم و $\mathbf{s}_{n\times 1}$ به عنوان بردار مجهول). اما در این‌جا شرط تنک‌بودن $\mathbf{s}_{n\times 1}$ مجموعه جواب‌ها را محدود می‌کند. سوال اصلی در این است که تحت چه شرایطی جواب به اندازه کافی تنک در محدوده جواب یکتا است و در صورت یکتابودن چگونه می‌توان با داشتن بردار $\mathbf{y}_{m\times 1}$ به بردار تنک $\mathbf{s}_{n\times 1}$ دست یافت. همان‌طور که در ابتدا اشاره شد، روش بازسازی از قبل معلوم بود: کمیته کردن نرم ℓ_1 (BP[۸]). این روش نه تنها در عمل کارآمد است بلکه در مقاله [۴۰] نشان داده شده که در حالت نظری نیز به روش بهینه بازسازی بسیار نزدیک است[۹]. پیش از آن که به بررسی شرایط لازم بر روی ماتریس $\mathbf{A}$ بپردازیم، روش BP را کمی دقیق‌تر مطالعه می‌کنیم. در مسأله بازسازی، به دنبال یافتن کمینه‌کننده عبارت زیر

[۵] Non-Adaptive
[۶] Sensing Matrix
[۷] Under-determined
[۸] Basis Pursuit
[۹] Near Optimal

هستیم:

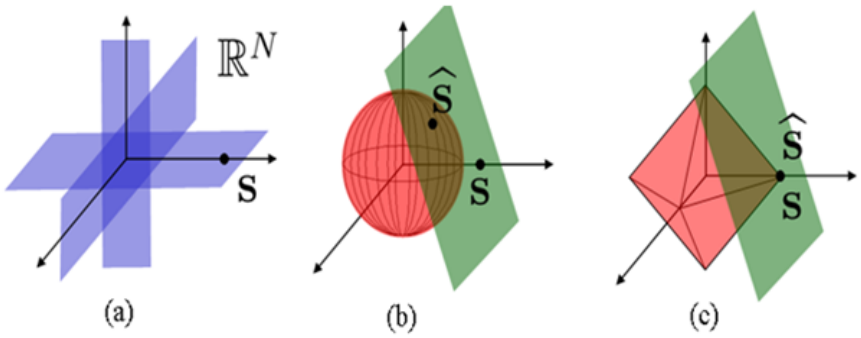
$$\arg\min_{\mathbf{s}} \|\mathbf{s}_{n\times 1}\|_{\ell_0} \quad s.t. \quad \mathbf{\Phi}_{m\times n}\cdot\mathbf{\Psi}_{n\times n}\cdot\mathbf{s}_{n\times 1}=\mathbf{y}_{m\times 1} \tag{۴-۱}$$

که منظور از $\|.\|_{\ell_0}$ نرم صفر بردار است (تعداد درایه‌های ناصفر). از آنجا که نرم صفر محدب نیست (کلا $\|.\|_{\ell_p}$ برای $0 \leq p < 1$ محدب نیست) و از آن مهم‌تر، مشتق‌پذیر نیست، مساله NP-Complete[۱۰] تلقی می‌شود و حل آن منوط به جستجوی کامل است که به وضوح در ابعاد بالا غیرعملی است. یکی از راهکارهای مهم در حل سیستم‌های خطی با فرض تنک‌بودن جواب (مشابه صورت مساله مطرح شده)، تقریب‌زدن نرم صفر با نرمی از مرتبه بالاتر است که قابلیت کمینه‌کردن آن ساده‌تر باشد. طبیعی است که هر چه نرم مرتبه پایین‌تری استفاده شود، جواب حاصل از کمینه‌سازی، به جواب حاصل از کمینه‌سازی نرم صفر نزدیک‌تر خواهد بود. از آنجا که نرم یک ($\|.\|_{\ell_1}$) نزدیک‌ترین نرم محدب به نرم صفر است، جایگزینی نرم صفر با نرم یک، منطقی‌ترین تقریب به نظر می‌رسد:

$$\arg\min_{\mathbf{s}} \|\mathbf{s}_{n\times 1}\|_{\ell_1} \quad s.t. \quad \mathbf{\Phi}_{m\times n}\cdot\mathbf{\Psi}_{n\times n}\cdot\mathbf{s}_{n\times 1}=\mathbf{y}_{m\times 1} \tag{۵-۱}$$

برای مساله اخیر راه‌حل‌هایی به کمک برنامه‌نویسی خطی[۱۱] معرفی شده‌اند که به خانواده Basis Pursuit معروفند و پیچیدگی محاسباتی آن‌ها از مرتبه n^3 است [۲۲، ۴۱، ۵۰، ۹۰]. به عبارت بهتر، با افزایش مرتبه نرم، پیچیدگی محاسباتی را کاهش داده‌ایم؛ حال آن‌که به احتمال زیاد جواب حاصل تا حدی با جواب اصلی متفاوت است. در صورتی که نرم صفر با نرم دو ($\|.\|_{\ell_2}$) تقریب زده شود، این روند مشهودتر خواهد شد: کمینه‌سازی براساس نرم دو جزء معروف‌ترین مسائل مهندسی بشمار می‌رود و راه‌حل‌های فراوانی برای آن پیشنهاد شده که پیچیدگی محاسباتی آن‌ها به مراتب کمتر از Basis Pursuit است. از جمله این روش‌ها می‌توان به استفاده از شبه-وارون[۱۲]، SD[۱۳]، CG[۱۴] و RLS[۱۵] اشاره کرد. روش رایج دیگر در بازسازی سیگنال‌های تنک Matching Pursuit است که به صورت تکراری و با روش Greedy جواب را تقریب می‌زند (پیچیدگی محاسباتی از مرتبه n) [۵۱، ۸۹، ۹۱]. نتایج شبیه‌سازی‌ها در کاربردهای مختلف حاکی از آن است که تقریب حاصل از نرم دو، در

[۱۰] Non-Polynomial Time
[۱۱] Linear Programming
[۱۲] Pseudo-Inverse
[۱۳] Steepest Descent
[۱۴] Conjugate Gradient
[۱۵] Recursive Least Squares



شکل ۱–۱: (a) یک بردار تنک در فضای سه بعدی، (b) تقریب حاصل از کمینه‌سازی نرم ℓ_2 برای این بردار، (c) تقریب حاصل از کمینه‌سازی نرم ℓ_1 برای این بردار [۱۰].

حالت کلی مناسب نیست در حالی که تقریب حاصل از نرم یک، هنگامی که تعداد نمونه‌ها کافی باشد، به سیگنال اصلی بسیار نزدیک است [۹۶] (با افزایش تعداد نمونه‌ها به اندازه زیاد، تمام نرم‌ها به یک جواب منجر می‌شوند، حال آن که تعداد نمونه‌های مورد نیاز در نرم‌های مختلف، متفاوت است؛ بطور اخص، نرم دو در قیاس با نرم یک، به تعداد نمونه‌های بسیار بیشتری نیاز دارد). شکل ۱–۱ مثالی از کمینه‌سازی با هر دو نرم یک و دو را نشان می‌دهد.

حال به سراغ ماتریس $\mathbf{A}$ می‌رویم. برای آن‌که بتوان هر سیگنال k–تنک را پس از نمونه‌برداری توسط $\mathbf{A}$ بازسازی کرد، باید هیچ دو بردار k–تنک متفاوتی نمونه‌های یکسانی تولید نکنند. در نتیجه تفاضل هیچ دو بردار k–تنکی (که در حالت کلی $2k$–تنک است) نباید در فضای پوچ این ماتریس قرار گیرد. پس شرط لازم برای بازسازی کامل آن است که هر $2k$ انتخاب از ستون‌های $\mathbf{A}$ مستقل خطی باشند. ماتریس‌های واندرموند[۱۶] $2k$ سطری از جمله ماتریس‌های معروفی هستند که چنین خاصیتی دارند. اما نکته منفی در این ماتریس‌ها، ناپایدار شدن آن‌ها در حالت حدی $n \to \infty$ است. یعنی قابلیت بازسازی منوط به دقت بسیار بالا در محاسبات و عدم حضور نویز جمعی است. برای یافتن چاره، باید پایداری جواب نسبت به نویز نیز لحاظ شود. یکی از ابزارهای قوی در مبحث نمونه‌برداری فشرده که نه تنها قابلیت بازسازی بلکه پایداری را نیز تضمین می‌کند، شرط RIP [۲۳] است.

می‌گوییم ماتریس $\mathbf{A}_{m\times n}$ شرط RIP مرتبه k را با ثابت δ_k ($0 \le \delta_k < 1$) ارضا می‌کند اگر برای هر بردار

[۱۶] Vandermonde

k–تنک مانند $\mathbf{s}_{n\times 1}$ داشته باشیم:

$$1-\delta_k \leq \frac{\|\mathbf{A}_{m\times n}\cdot \mathbf{s}_{n\times 1}\|_{\ell_2}^2}{\|\mathbf{s}_{n\times 1}\|_{\ell_2}^2} \leq 1+\delta_k \qquad (۶–۱)$$

به عبارت بهتر، نه تنها هیچ یک از بردارهای k–تنک در فضای پوچ ماتریس $\mathbf{A}$ قرار نمی‌گیرند، بلکه فاصله تضمین‌شده‌ای (δ_k) را نسبت به این فضا حفظ می‌کنند.

با وجود تمام نکات مثبت در مورد شرط RIP، از نظر محاسباتی بررسی این که یک ماتریس داده شده شرط RIP از چه مرتبه‌ای را ارضا می‌کند، NP-Hard است. یعنی جز در موارد خاصی نمی‌توان برقراری یا عدم برقراری شرط RIP را در یک ماتریس اثبات کرد. در [۲۵] با کمک گرفتن از ماتریس‌های تصادفی، وجود ماتریس‌های $\mathbf{A}_{m\times n}$ با $m \geq \mathcal{O}\big(k\log\frac{n}{k}\big)$ که شرط RIP مرتبه k را با ثابت دلخواه δ_k ارضا می‌کنند، اثبات شده است، اما تاکنون هیچ روش ساختاری برای تهیه چنین ماتریس‌هایی ارائه نشده است.

ضریب همدوسی ماتریس $\mathbf{A}$ که به صورت زیر تعریف می‌شود، در طراحی ماتریس‌های حسگر اهمیت بسیاری دارد:

$$\mu_{\mathbf{A}} \triangleq \max_{i\neq j} \frac{|\langle \mathbf{a}_i, \mathbf{a}_j\rangle|}{\|\mathbf{a}_i\|\cdot\|\mathbf{a}_j\|}, \qquad (۷–۱)$$

که $\mathbf{a}_i$ و $\mathbf{a}_j$ ستون‌های متمایز ماتریس $\mathbf{A}$ هستند. هرچه ضریب همدوسی ماتریس کوچکتر باشد، ماتریس به متعامدبودن نزدیک‌تر است. در فصل آینده نشان می‌دهیم که کوچک‌بودن ضریب همدوسی ماتریس، شرط RIP را تضمین می‌کند و در نتیجه شرط قوی‌تری نسبت به RIP به‌شمار می‌رود. همچنین محاسبه ضریب همدوسی یک ماتریس از نظر محاسباتی به صورت $\mathcal{O}(n^2)$ با افزایش n رشد می‌کند که در موارد عملی قابل قبول است. براساس یک نامساوی معروف که به نام Welch شهرت دارد [۸۷] در ماتریس‌های $m\times n$ که $m<n$

$$\mu_{\mathbf{A}} \geq \sqrt{\frac{n}{m(n-m)}} \qquad (۸–۱)$$

رابطه فوق نشان می‌دهد زمانی که $m \ll n$، $\mu_{\mathbf{A}}$ را نمی‌توان خیلی کوچکتر از $\frac{1}{\sqrt{m}}$ اختیار کرد. در نتیجه اگر بخواهیم با استفاده از ضریب همدوسی شرط RIP را ارضا کنیم، همواره $m \geq \mathcal{O}(k^2)$ و هیچ‌گاه به حالت پیش‌بینی شده $m=\mathcal{O}\big(k\log\frac{n}{k}\big)$ نخواهیم رسید. با این حال استفاده از ضریب همدوسی در طراحی ماتریس حسگر تاکنون تنها ابزار مناسب بوده است.

۱-۳ کاربردهای نمونه‌برداری فشرده

در پنج سال اخیر، به دلیل استقبال شدید محققین از مبحث جدید نمونه‌برداری فشرده، گام‌های موثر و ارزنده‌ای در این راستا برداشته شده است. به طور خاص، پیشرفت‌های چشمگیری در زمینه روش‌های بازسازی سیگنال تنک از روی نمونه‌های فشرده حاصل شده است. با وجود آن که مرحله بازسازی قسمتی مهم از نمونه‌برداری فشرده را در بر دارد، کمینه‌سازی نرم ℓ_1 از این مبحث قدمت بیشتری دارد و نمی‌توان آن را جزء دستاوردهای این زمینه تلقی کرد. حال آن که کاربرد این روش‌های بازسازی در زمینه‌های دیگر همچون بازیابی عکس‌های محو شده [۲۹]، جداسازی منابع [۱۴]، رادار [۴۴] و حتی آشکارسازی نور ستارگان [۱۶] منجر به بهبود و ارتقاء روش‌های کلاسیک شده است. اغلب استفاده از روش‌های بازسازی سیگنال‌های تنک در زمینه‌های دیگر را به اشتباه جزء کاربردهای مبحث نمونه‌برداری فشرده به شمار می‌آورند. با وجود آن که شکوفایی این مبحث موجب مطالعه و پیشرفت این روش‌ها شده است، نمی‌توان روش‌های بازسازی را جزء کاربردهای این زمینه به حساب آورد.

از ابتدا، ایجاد تحول در صنعت عکاسی و مبدل‌های آنالوگ به دیجیتال جزء اهداف اصلی نمونه‌برداری فشرده بوده‌اند که در مطالعات نظری امری میسر به نظر می‌رسد ولی تاکنون چنین طرح‌هایی جامه عمل به خود ندیده‌اند و کماکان جزء اهداف اصلی این مبحث به شمار می‌روند. نقطه ضعف‌های اصلی این مبحث در رسیدن به این اهداف عبارتند از : ۱) پایه‌ریزی عمده مطالب بر اساس سیگنال‌های گسسته و نه پیوسته ؛ ۲) استفاده از ساختارهای تصادفی.

از جمله کاربردهای نمونه‌برداری فشرده که محقق شده است، عکس‌برداری به روش تشدید مغناطیسی (MRI) است [۶۸]. در این روش پرتوهای اشعه X از جسمی (مثل بدن انسان) عبور داده می‌شوند و به دلیل خواص اپتیکی محیط، تبدیل فوریه فضایی جسم در راستای پرتو تابانده شده، دریافت می‌شود. به این صورت، به تعداد پرتوهای تابانده شده، ضریب فوریه از جسم به دست می‌آید و در نهایت براساس ضرایب فوریه بدست آمده، محتوی سه بعدی جسم تقریب زده می‌شود. از آن‌جا که اجسام معمولی و به ویژه بافت‌های بدن، از قسمت‌های تکه‌ای یکنواخت تشکیل شده‌اند، مشتق مکانی جسم تنک است (مشتق مکانی لبه‌ها را آشکار می‌کند). از طرف دیگر، نمونه‌ها به طور ناقص از تبدیل فوریه جسم انتخاب شده‌اند. به دلیل مضربودن پرتوهای اشعه X، مطلوب است که تعداد پرتوهای تابانده شده تا حد ممکن کاهش یابد، به عبارت بهتر، تمام شرایط

نمونه‌برداری فشرده در این‌جا وجود دارد.

مثال دیگری از کاربردهای نمونه‌برداری فشرده، تخمین کانال در ارتباطات OFDM است [۸۵، ۸۶] . در روش OFDM، اطلاعات ارسالی بر روی حامل‌های فرکانسی مجزا و هم‌فاصله قرار داده می‌شوند. برای قابلیت تخمین کانال، بر روی تعدادی از حامل‌ها داده‌های از پیش تعیین‌شده قرار می‌گیرد که گیرنده از قبل نسبت به آن‌ها واقف است. پس از عبور داده از کانال، دنباله ارسالی و پاسخ ضربه کانال به هم پیچیده[۱۷] می‌شوند که معادل ضرب در حوزه فرکانس است؛ در نتیجه حامل‌های فرکانسی به طور مستقل تحت تاثیر کانال قرار می‌گیرند. بنابراین گیرنده قادر است که پاسخ فرکانسی کانال را در فرکانس‌های از پیش تعیین‌شده با دقت مناسبی تخمین زند. برای تخمین کل کانال، کافی است به این نکته توجه شود که پاسخ ضربه کانال معمولاً تنک است. بنابراین با داشتن تعدادی از ضرایب فوریه یک سیگنال تنک (که در انتخاب ضرایب نسبتاً آزاد هستیم) به دنبال تخمین آن هستیم. به وضوح این مسأله حالت خاصی از نمونه‌برداری فشرده است.

با توجه به نتایج نظری بدست آمده و قابلیت‌های فناوری، حدس زده می‌شود که در آینده نزدیک از نمونه‌برداری فشرده در شبکه حسگرها [۱۸] و ضبط سیگنال‌های حیاتی توان پایین به طور عملی استفاده شود.

۱-۴ تکمیل ماتریس

پس از اوج گرفتن مبحث نمونه‌برداری فشرده، تعمیم‌های آن به ابعاد بالاتر نیز مطرح شد. مبحث تکمیل ماتریس[۱۹] یکی از این موارد است. هنگامی که به جای یک بردار، با یک ماتریس روبرو هستیم، شرط تنک‌بودن معادل با کوچک‌بودن رتبه ماتریس است. در حقیقت هنگامی که ماتریس به عنوان یک عملگر در نظر گرفته شود، مقادیر ویژه و یا مقادیر تکین آن اهمیت بسیار بیشتری نسبت به درایه‌ها پیدا می‌کنند. بنابراین، تنک‌بودن یک ماتریس بیشتر معادل با تنک‌بودن مقادیر ویژه آن است تا المانهایش. علاوه بر این، وجود تعداد زیادی صفر در بین مقادیر ویژه به معنی پایین‌بودن رتبه ماتریس نسبت به ابعادش است.

در مبحث تکمیل ماتریس، فرض بر این است که تعدادی از درایه‌های یک ماتریس کم‌رتبه در اختیار است و هدف کامل‌کردن بقیه درایه‌هاست به نحوی که رتبه ماتریس حداقل شود. در [۲۱] نشان داده شده که نرم

[۱۷] convolve

[۱۸] Sensor Networks

[۱۹] Matrix Completion

هسته‌ای[۲۰] ماتریس نقش نرم ℓ_1 در بردارها را ایفا می‌کند و روش کمینه‌کردن نرم هسته‌ای خواصی مشابه با روش BP دارد. یکی از کاربردهای تکمیل ماتریس که در [۲۱] ذکر شده است، مسأله داوری در فستیوال فیلم است. فرض کنید ۱۰۰ فیلم در مسابقه شرکت داده شده‌اند و تنها ۱۰ داور برای بررسی این فیلم‌ها در نظر گرفته‌شده‌است؛ به دلیل وقت کم و تعداد زیاد فیلم‌ها، هر فیلم تنها توسط ۴ یا ۵ داور مورد قضاوت قرار می‌گیرد. اکنون اگر ماتریس ۱۰۰×۱۰ مربوط به نمرات داوران در مورد فیلم‌ها را تشکیل دهیم، بسیاری از درایه‌ها در دسترس نیستند. اما به طور شهودی براساس نظرات یک داور و مقایسه آن با نظرات داورهای دیگر، می‌توان بقیه درایه‌های ماتریس را تخمین زد. فرض کنید تنها ۴ عامل فیلمنامه، کارگردانی، بازی هنرپیشه‌ها و جلوه‌های ویژه از نظر داورها اهمیت دارد و هر داور براساس علاقه شخصی خود به نوعی بین این عوامل وزن‌دهی می‌کند و به هر فیلم امتیاز می‌دهد. در این صورت رتبه ماتریس ۱۰۰×۱۰ برابر با ۴ خواهد بود و برای تخمین درایه‌های نامعلوم ماتریس می‌توان از روش‌های تکمیل ماتریس سود جست.

۱-۵ طبقه بندی مطالب پایان‌نامه

در این فصل (مقدمه) به توضیح مقدمات و تعاریف مورد استفاده در نمونه‌برداری فشرده پرداختیم. در فصل ۲ کلیات نمونه‌برداری فشرده از قبیل پیش زمینه‌ها و دیدگاه‌های مختلف نسبت به آن توضیح داده خواهد شد. از آن‌جا که نوآوری این پایان‌نامه بیشتر در طراحی روش‌های نمونه‌برداری است، در فصل ۳ به مرور روش‌های کنونی در نمونه‌برداری فشرده می‌پردازیم. ماتریس‌های حسگر جدید که در این پایان‌نامه معرفی شده‌اند در فصل ۴ مطرح می‌شوند. پس از معرفی روش‌های خطی، در فصل ۵ به روش‌های غیرخطی و مزایایی که می‌توانند داشته باشند می‌پردازیم. هدف اصلی در نمونه‌برداری فشرده متحول‌کردن روش‌های کنونی نمونه‌برداری و فشرده‌سازی است که در عمل بر روی داده‌های پیوسته صورت می‌گیرند. برای ورود به مبحث سیگنال‌های پیوسته لازم است که تعاریف مقدماتی از قبیل تنک‌بودن و فشرده‌پذیری مجدداً برای این سیگنال‌ها مطالعه شود که این امر در فصل ۶ صورت گرفته است. ماتریس‌ها تعمیم‌های طبیعی بردارها به شمار می‌روند و همان‌طور که در قسمت قبل اشاره شد، مفهوم تنک‌بودن برای ماتریس‌ها به صورت کم‌بودن رتبه مطرح می‌شود. در فصل ۷ به کمک عملگرهای غیرخطی به دنبال کاهش رتبه ماتریس‌ها هستیم. در انتها با یک جمع‌بندی در فصل ۸ ، پایان‌نامه را به اتمام می‌رسانیم.

[۲۰] Nuclear Norm

برای تفکیک نوآوری‌ها در این پایان‌نامه وکارهای موجود، این مطالب در فصل‌های مجزا گنجانده شده‌اند. به طور خاص فصل‌های ۴، ۵، ۶ و ۷ نتایج این پایان‌نامه هستند که در مقالات [۲، ۳، ۴، ۵، ۶، ۷، ۸] ارایه شده‌اند.

فصل ۲

کلیات نمونه‌برداری فشرده

۲-۱ مقدمه

در این فصل نمونه‌برداری فشرده را از چند دیدگاه بررسی می‌کنیم. ابتدا پیش زمینه‌های ریاضی که باعث پیدایش نظریه نمونه‌برداری فشرده شده است را بررسی می‌کنیم. عرض گلفاند مجموعه‌ها مبحثی است که سال‌ها قبل از نمونه‌برداری فشرده مطرح شده است و ارتباط تنگاتنگی با این نظریه دارد [۳۸]. از طرف دیگر، نمونه‌برداری فشرده از دیدگاه نظریه‌ی اطلاعات یک روش فشرده‌سازی تلقی می‌شود. بدین منظور نمونه‌برداری فشرده از دیدگاه نرخ اعوجاج[۱] را نیز بررسی می‌کنیم و در ادامه مرور مختصری بر روش‌های بازسازی ارائه می‌دهیم.

۲-۲ عرض گلفاند

فرض کنید S زیرمجموعه‌ای فشرده از $\mathbb{R}^n$ و m عددی طبیعی باشد. عرض گلفاند[۲] مجموعه S را به صورت زیر تعریف می‌کنیم:

$$p \geq ۱ : \quad d^m(S)_{\ell_p} = \inf_{Y} \ \sup\{\|\mathbf{x}\|_{\ell_p} \mid \mathbf{x} \in S \cap Y\} \tag{۱-۲}$$

که در آن infimum روی تمام زیرفضاهای Y ($\mathbb{R}^n$) با بعد حداقل $n-m$ گرفته می‌شود. نشان می‌دهیم که عملکرد نمونه‌برداری فشرده روی مجموعه S، رابطه بسیار نزدیکی با عرض گلفاند مجموعه دارد [۳۳]. فرض

[۱] Rate-Distortion

[۲] Gelfand's width

کنید S زیرمجموعه‌ای کراندار (با نرم ℓ_p) از $\mathbb{R}^n$ باشد به طوری که قرینه جمعی هر بردار داخل S عضوی از S باشد:

$$\forall\, s \in S: \quad -s \in S \tag{۲-۲}$$

همچنین فرض کنید جمع هر دو عضو S ، ضریبی (ثابت) از یک عنصر S باشد:

$$\exists\, c_\circ \in \mathbb{R}: \quad S + S \subset c_\circ S \tag{۳-۲}$$

مثلاً اگر S مجموعه بردارهای با نرم کمتر از واحد باشد، می‌توان $c_\circ = ۲$ اختیار کرد. می‌خواهیم اعضای S را توسط ماتریس $\mathbf{\Phi}_{m\times n}$ نمونه‌برداری کنیم (در حالت کلی اعضای $\mathbb{R}^n$ را نمونه‌برداری می‌کنیم اما در بازسازی به دنبال یافتن ورودی مناسب در S هستیم). برای بازسازی نیز از تابعی به نام D استفاده می‌کنیم که لزوماً خطی نیست:

$$\begin{cases} \mathbf{y}_{m\times ۱} = \mathbf{\Phi}_{m\times n} \cdot \mathbf{x}_{n\times ۱} \\ \hat{\mathbf{x}}_{n\times ۱} = D\big(\mathbf{y}_{m\times ۱}\big) \end{cases} \tag{۴-۲}$$

از آنجا که معیار سنجش ما نرم ℓ_p است، خطای حاصل از بازسازی توسط D به صورت زیر خواهد بود:

$$E\big(\mathbf{x}_{n\times ۱}, \mathbf{\Phi}_{m\times n}, D\big) = \big\|\mathbf{x}_{n\times ۱} - \hat{\mathbf{x}}_{n\times ۱}\big\|_{\ell_p} = \big\|\mathbf{x}_{n\times ۱} - D\big(\mathbf{\Phi}_{m\times n}\cdot \mathbf{x}_{n\times ۱}\big)\big\|_{\ell_p} \tag{۵-۲}$$

برای بررسی عملکرد نمونه‌برداری-بازسازی تولید شده توسط $\mathbf{\Phi}_{m\times n}$ وD، تنها خطای ایجاد شده روی یک عضو کفایت نمی‌کند بلکه باید به نوعی خطای ایجاد شده روی تمام حالات ورودی بررسی شود. مثلاً:

$$E\big(S, \mathbf{\Phi}_{m\times n}, D\big)_{\ell_p} = \sup_{\mathbf{x}_{n\times ۱}\in S} E\big(\mathbf{x}_{n\times ۱}, \mathbf{\Phi}_{m\times n}, D\big)_{\ell_p} \tag{۶-۲}$$

در نمونه‌برداری فشرده، در پی بهترین زوج D و $\mathbf{\Phi}_{m\times n}$ هستیم؛ به عبارت دیگر، می‌خواهیم ماتریس حسگر و روش بازسازی به نوعی انتخاب شوند که خطای ایجاد شده روی اعضایS به حداقل ممکن برسد. با انتخاب بهترین زوج D و $\mathbf{\Phi}_{m\times n}$ (با ثابت نگه داشتن m) $E\big(\mathbf{x}_{n\times ۱}, \mathbf{\Phi}_{m\times n}, D\big)_{\ell_p}$ به کمترین مقدار خود می‌رسد که این مقدار، بهترین سطح خطای ایجاد شده در بازسازی اعضای S است و فقط تابعی از S و m (و نرم ℓ_p) خواهد بود:

$$E_m(S)_{\ell_p} = \inf_{\mathbf{\Phi}_{m\times n}, D} E\big(S, \mathbf{\Phi}_{m\times n}, D\big)_{\ell_p} \tag{۷-۲}$$

اکنون می‌توان نشان داد که $E_m(S)_{\ell_p}$ به طور جالبی به $d^m(S)_{\ell_p}$ مربوط است : از آنجا که ماتریس $\mathbf{\Phi}_{m\times n}$ در نمونه‌برداری فشرده یک ماتریس مستطیلی افقی است ($m \le n$)، بعد فضای پوچ ($dim\ \mathcal{N}_{\mathbf{\Phi}}$) آن حداقل $n-m$ خواهد بود که انتخاب مناسبی برای Y در تعریف عرض گلفاند است:

$$d^m(S)_{\ell_p} = \inf_{Y} \sup_{\mathbf{x}\in S\cap Y} \|\mathbf{x}\|_{\ell_p} \le \sup_{\mathbf{x}\in S\cap N_{\mathbf{\Phi}}} \|\mathbf{x}\|_{\ell_p} \tag{۸-۲}$$

از طرفی، برای هر $\mathbf{x} \in S\cap\mathcal{N}_{\mathbf{\Phi}}$ داریم: $\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1} = 0_{m\times 1}$ یعنی برای این بردارهای $\mathbf{x}$، $D(\mathbf{\Phi},\mathbf{x}) = D(0)$ و ثابت است.

$$\mathbf{x}\in S\cap\mathcal{N}_{\mathbf{\Phi}} \Rightarrow -\mathbf{x}\in S\cap\mathcal{N}_{\mathbf{\Phi}} \Rightarrow D(\mathbf{\Phi}\cdot\mathbf{x}) = D(-\mathbf{\Phi}\cdot\mathbf{x}) = D(0)$$

$$\begin{aligned} &\|\mathbf{x}-D(0)\|_{\ell_p} + \|-\mathbf{x}-D(0)\|_{\ell_p} \ge 2\|\mathbf{x}\|_{\ell_p} \\ &\Rightarrow \begin{cases} E(\mathbf{x}_{n\times 1},\mathbf{\Phi}_{m\times n},D)_{\ell_p} \ge \|\mathbf{x}\|_{\ell_p} \\ \qquad or \\ E(-\mathbf{x}_{n\times 1},\mathbf{\Phi}_{m\times n},D)_{\ell_p} \ge \|\mathbf{x}\|_{\ell_p} \end{cases} \end{aligned} \tag{۹-۲}$$

در نتیجه برای هر $\mathbf{x}\in S\cap\mathcal{N}_{\mathbf{\Phi}}$ عنصری در این مجموعه وجود دارد که خطای بازسازی آن حداقل به اندازه $\|\mathbf{x}\|_{\ell_p}$ است:

$$\begin{aligned} E(S,\mathbf{\Phi}_{m\times n},D)_{\ell_p} &= \sup_{\mathbf{x}_{n\times 1}\in S} E(\mathbf{x}_{n\times 1},\mathbf{\Phi}_{m\times n},D)_{\ell_p} \\ &\ge \sup_{\mathbf{x}\in S\cap\mathcal{N}_{\mathbf{\Phi}}} E(\mathbf{x}_{n\times 1},\mathbf{\Phi}_{m\times n},D)_{\ell_p} \ge \sup_{\mathbf{x}\in S\cap\mathcal{N}_{\mathbf{\Phi}}} \|\mathbf{x}\|_{\ell_p} \end{aligned} \tag{۱۰-۲}$$

پس مشاهده شد که عرض گلفاند مجموعه، یک کران پایین برای خطای بازسازی در هر انتخاب $D,\mathbf{\Phi}_{m\times n}$ است. پس:

$$E_m(S)_{\ell_p} \ge d^m(S)_{\ell_p} \tag{۱۱-۲}$$

نامساوی فوق به تنهایی اهمیت خاصی ندارد زیرا فقط یک کران پایین است و معلوم نیست که تا چه حد میتوان به این کران نزدیک شد. نکته قابل توجه این است که $E_m(S)_{\ell_p}$ از بالا نیز محدود به ضریبی از $d^m(S)_{\ell_p}$ می‌شود: فرض کنید Y زیرفضایی با بعد $n-m$ از $\mathbb{R}^n$ باشد؛ در نتیجه بعد فضای متعامد Y ($Y^\perp$) برابر با m خواهد بود. فرض کنید $\mathbf{v}_1,\dots,\mathbf{v}_m$ پایه‌هایی برای این زیر فضا ($Y^\perp$) باشند. اکنون ماتریس $\mathbf{\Phi}$ نمونه‌برداری را بر حسب این

Y چنین می‌سازیم:

$$\mathbf{\Phi}_{m\times n}=\begin{bmatrix}\mathbf{v}_1^T\\ \vdots\\ \mathbf{v}_m^T\end{bmatrix} \tag{۱۲-۲}$$

همچنین تابع بازسازی D را چنین تعریف می‌کنیم: برای یک $\mathbf{u}_{m\times 1}\in\mathbb{R}^m$ دلخواه، اگر $\mathbf{u}_{m\times 1}\in\mathbf{\Phi}\cdot S$، $D(\mathbf{u}_{m\times 1})$ را به دلخواه یکی از اعضای S مانند $\mathbf{a}$ تعریف می‌کنیم که $\mathbf{u}_{m\times 1}=\mathbf{\Phi}\cdot\mathbf{a}$؛ درغیر این صورت $D(\mathbf{u}_{m\times 1})$ را بـه طور تصادفی یکی از اعضای S قرار می‌دهیم. با مفروضات فوق، $E_m(S)_{\ell_p}$ را بدست می‌آوریم. اگر $\mathbf{x}_{n\times 1}\in S$، داریم:

$$\begin{aligned}\mathbf{\Phi}_{m\times n}\big(\mathbf{x}_{n\times 1}-D(\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1})\big) &= \mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1}-\mathbf{\Phi}_{m\times n}\cdot D(\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1})\\ &= \mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1}-\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1}=0_{m\times 1}\\ \Rightarrow\quad & \mathbf{x}_{n\times 1}-D(\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1})\in\mathcal{N}_{\mathbf{\Phi}_{m\times n}}\\ \Rightarrow\quad & \mathbf{x}_{n\times 1}-D(\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1})\in Y\end{aligned} \tag{۱۳-۲}$$

از طرفی طبق فرض اولیه، تفاضل هر دو عضو S ضریبی از یک عضو S خواهد بود:

$$\exists c_0\in\mathbb{R}:\ \frac{\mathbf{x}_{n\times 1}-D(\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1})}{c_0}\in S\Rightarrow\frac{\mathbf{x}_{n\times 1}-D(\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1})}{c_0}\in S\cap Y \tag{۱۴-۲}$$

حال داریم :

$$\begin{aligned}&E(S,\mathbf{\Phi}_{m\times n},D)_{\ell_p}=c_0\sup_{\mathbf{x}_{n\times 1}\in S}\left\|\frac{\mathbf{x}_{n\times 1}-D(\mathbf{\Phi}_{m\times n}\cdot\mathbf{x}_{n\times 1})}{c_0}\right\|_{\ell_p}\le c_0\sup_{\mathbf{x}_{n\times 1}\in S\cap Y}\|\mathbf{x}_{n\times 1}\|_{\ell_p}\\ &\Rightarrow\inf_{\mathbf{\Phi}_{m\times n},D}E(S,\mathbf{\Phi}_{m\times n},D)_{\ell_p}\le c_0\inf_{Y}\sup_{\mathbf{x}_{n\times 1}\in S\cap Y}\|\mathbf{x}_{n\times 1}\|_{\ell_p}\\ &\Rightarrow E_m(S)_{\ell_p}\le c_0\ d^m(S)_{\ell_p}\end{aligned} \tag{۱۵-۲}$$

پس در کل خواهیم داشت:

$$d^m(S)_{\ell_p}\le E_m(S)_{\ell_p}\le c_0\ d^m(S)_{\ell_p} \tag{۱۶-۲}$$

در نتیجه، یافتن عملکرد بهترین نحوه نمونه‌برداری–بازسازی روی S تقریباً معادل با یافتن عرض گلفاند این مجموعه است. متاسفانه یافتن عرض گلفاند مجموعه‌ها جز در حالاتی خاص (و آن هم نه به سادگی) هنوز

مساله‌ای باز محسوب می‌شود [۶۳]. از موارد خاص می‌توان به عرض گلفاند گوی واحد n بعدی با نرم ℓ_p اشاره کرد ($U(\ell_p^n)$) :

$$U(\ell_p^n) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_{\ell_p} = ۱\} \tag{۲–۱۷}$$

در چنین حالتی $d^m\big(U(\ell_p^n)\big)_{\ell_q}$ ، بین دو کران بالایی و پایینی محدود می‌شود که کران بالایی ضریبی از کران پایینی است اما مقدار دقیقی برای این ضریب بدست نیامده است؛ تنها مرتبه بزرگی $d^m\big(U(\ell_p^n)\big)_{\ell_q}$ مشخص است [۶۳]. به هر حال، عرض گلفاند را می‌توان به عنوان مرجع سنجشی برای خطای ایجاد شده در یک روش نمونه‌برداری فشرده تلقی کرد (مانند باند Cramer-Rao در تخمین؛ با این تفاوت که در کمترین خطا، لزوماً به عرض گلفاند دست نخواهیم یافت).

۲–۳ بررسی نمونه‌برداری فشرده از دیدگاه نظریه نرخ–اعوجاج

در نمونه‌برداری فشرده، هـدف ذخیره‌سازی دسته خاصی از سیگنال‌ها به کمک تعداد نمونه کم است. پس می‌توان به نوعی این روش را جزء روش‌های کدگذاری منبع طبقه بندی کرد. از آنجا که این روش لزوماً به بازسازی دقیق سیگنال منجر نمی‌شود، یک روش فشرده‌سازی با اتلاف محسوب می‌شود؛ در نتیجه بررسی عملکرد آن باید توسط ابزارهای مطرح شده در نظریه نرخ–اعوجاج[۳] صورت پذیرد. فرض کنید معیار سنجش اعوجاج ($\mathcal{D}$) به صورت MSE باشد :

$$\mathcal{D} = \mathscr{E}\{(x - \hat{x})^۲\} \tag{۲–۱۸}$$

همچنین فرض کنید نرخ ذخیره‌سازی اطلاعات بر حسب بیتR باشد؛ مثلاً اگر سیگنالی را که در بازه‌ای به طول c توزیع یکنواخت دارد، با استفاده از چندی‌سازی[۴] با پله‌های Δ گسسته کنیم :

$$R = \log_۲\left(\frac{c}{\Delta}\right) \tag{۲–۱۹}$$

اگر توزیع ورودی را مشابه فرض فوق یکنواخت فرض کنیم، متغیر $x - \hat{x}$ (اختلاف مقدار اصلی با مقدار چندی‌شده) نیز در بازه $[-\frac{\Delta}{۲}, \frac{\Delta}{۲}-]$ توزیع یکنواخت خواهد داشت:

$$\mathcal{D} = \mathscr{E}\{(x - \hat{x})^۲\} = \frac{\Delta^۲}{۱۲} \tag{۲–۲۰}$$

[۳] Rate-Distortion

[۴] quantization

و در نتیجه، رابطه‌ی بین نرخ ذخیره‌سازی بیت و اعوجاج به صورت $\mathcal{D} = \frac{c^۲}{۱۲}۲^{-۲R}$ خواهد بود. در حالت کلی تر، اگر سیگنال ورودی توزیع دلخواه داشته باشد، در [۵۳] نشان داده شده است که بهترین (کمترین) میزان اعوجاج در چندی‌کردن این سیگنال با نرخ ثابت R در رابطه‌ی زیر صدق می‌کند:

$$\frac{۲^{۲h}}{۲\pi e}۲^{-۲R} \leq \mathcal{D} \leq \sigma^۲ ۲^{-۲R} \qquad (۲۱-۲)$$

که در آن h و $\sigma^۲$ به ترتیب آنتروپی پیوسته و واریانس سیگنال ورودی هستند. جمله‌ی اصلی در روابط اعوجاج که از نرخ حاصل می‌شود، $۲^{-۲R}$ است؛ یعنی وابستگی اعوجاج به نرخ بیت به صورت ضریبی از $۲^{-۲R}$ خواهد بود. اکنون مسأله‌ی نمونه‌برداری فشرده را برای تحلیل نرخ–اعوجاج مدل می‌کنیم: فرض کنید ماتریس $\mathbf{\Psi}_{n\times n}$ یک ماتریس متعامد–یکه تصادفی است که تنها در گیرنده (عملیات بازسازی) معلوم است و همچنین فرض کنید بردار تنک $\mathbf{s}_{n\times ۱}$ از kدرایه ناصفر تشکیل شده باشد که انتخاب این k محل با توزیع یکنواخت بین $\binom{n}{k}$ حالت مختلف صورت گرفته است و درایه‌های ناصفر به طور مستقل از هم، توزیع گوسی با متوسط صفر و واریانس یک دارند. بردار زمانی نیز به صورت $\mathbf{x}_{n\times ۱} = \mathbf{\Psi}_{n\times n} \cdot \mathbf{s}_{n\times ۱}$ در نظر گرفته می‌شود که پس از نمونه‌برداری توسط ماتریس $\mathbf{\Phi}_{m\times n}$ به $\mathbf{y}_{m\times ۱} = \mathbf{\Phi}_{m\times n} \cdot \mathbf{x}_{n\times ۱}$ تبدیل می‌شود. اکنون باید نمونه‌های بدست آمده (y_i) را چندی‌کرد:

$$\begin{aligned} \hat{y}_i &= Q(y_i) \\ \beta = \frac{\mathscr{E}\{|y_i - \hat{y}_i|^۲\}}{\mathscr{E}\{|y_i|^۲\}} &\quad , \quad \rho = ۱ - \beta \\ \Rightarrow \hat{y}_i &= \rho y_i + \nu_i \end{aligned} \qquad (۲۲-۲)$$

که در آن ν_i نویز چندی‌کردن است و واریانس آن برابر با $\beta(۱-\beta)\mathscr{E}\{|y_i|^۲\}$ است (ν_i با y_i نابسته خواهد بود اما $y_i - \hat{y}_i$ نسبت به y_i نابسته نیست):

$$\hat{\mathbf{y}} = \underbrace{\rho\mathbf{\Phi} \cdot \mathbf{\Psi}}_{\mathbf{A}} \cdot \mathbf{s} + \nu = \mathbf{A} \cdot \mathbf{s} + \nu \qquad (۲۳-۲)$$

فرض کنیم برای ذخیره هر نمونه ناصفر، R بیت اختصاص یافته باشد؛ در این صورت برای ذخیره‌سازی بردار $\mathbf{s}$ و در نتیجه $\mathbf{x}$، در کل Rk بیت اختصاص یافته‌است. اکنون درپی محاسبه اعوجاج هستیم. اگر از ابتدا تمام نمونه‌های $\mathbf{x}$ را بدون توجه به شرط تنک‌بودن $\mathbf{s}$ چندی می‌کردیم، به هر درایه $\frac{kR}{n}$ بیت اختصاص می‌یافت و در نتیجه:

$$\mathcal{D}_{direct} \propto ۲^{-۲\frac{kR}{n}} \qquad (۲۴-۲)$$

اگر می‌شد فقط نمونه‌های ناصفر $\mathbf{s}$ را ذخیره کرد، به غیر از k نمونه ناصفر، باید مکان‌های ناصفر نیز ذخیره می‌شد؛ پس بطور متوسط $R_۰ = \frac{\log_۲ \binom{n}{k}}{k}$ بیت از بیت‌های اختصاص داده شده به هر نمونه، باید به این امر اختصاص یابد. بنابراین:

$$\mathcal{D}_{adaptive} \propto ۲^{-۲(R-R_۰)} \tag{۲۵-۲}$$

در نمونه‌برداری فشرده که m نمونه در اختیار داریم، به هر نمونه $\frac{kR}{m}$ بیت اختصاص خواهد یافت. بدیهی است که با افزایش m (تعداد نمونه‌ها)، بازسازی از روی نمونه‌های چندسطحی نشده دقیق‌تر می‌شود؛ حال آنکه با افزایش m، هر نمونه با کیفیت پایین‌تری ذخیره می‌شود (نرخ کلی ثابت) و در نهایت اعوجاج سیگنال نهایی افزایش می‌یابد. در [۴۷] نشان داده شده که در نرخ‌های R متوسط به بالا، مقدار بهینه m از رابطه زیر بدست می‌آید:

$$m \approx \frac{۲kR\ln ۲}{۲R\ln ۲ - ۱} \tag{۲۶-۲}$$

برای m رابطه فوق چنین خواهیم داشت:

$$\mathcal{D}_{CS} \propto ۲^{-۲(R-R^*)} \tag{۲۷-۲}$$

که :

$$R^* = \frac{\log_۲(۲eR\ln ۲)}{۲} \tag{۲۸-۲}$$

مقایسه $\mathcal{D}_{CS}$ با $\mathcal{D}_{adaptive}$ حاکی از آن است که کیفیت بطور قابل ملاحظه‌ای افت کرده که تقریباً نتیجه واضحی است زیرا $\mathcal{D}_{adaptive}$ با تقریب خوبی بهترین کیفیت قابل دسترسی است. در بازسازی، ابتدا باید محل نمونه‌های ناصفر $\mathbf{s}$ تخمین زده شود و سپس مقادیر آن‌ها بدست آید؛ نکته منفی که در [۵۲] ذکر شده این است که حتی با افزایش تعداد نمونه‌ها (m) تضمینی برای تخمین درست از این مکان‌ها بدست نمی‌آید.

۲-۴ ارتباط نمونه‌برداری فشرده با نظریه کدگذاری

نویز جمعی مشکلی غیرقابل پرهیز در انواع مخابرات محسوب می‌شود. در مخابرات دیجیتال، به دلیل وجود نویز، تعدادی از بیت‌های ارسالی دستخوش تغییر می‌شوند که در صورت تصحیح نشدن، می‌توانند در گیرنده

مشکلات بسیاری بوجود آورند. سازوکار رایج در آشکارسازی و تصحیح بیت‌های خطا، استفاده از کدگذاری است. در روش کدگذاری قالبی[۵]، یک قالب kتایی از بیت‌های داده، با افزودن تعدادی بیت توازن[۶] به یک قالب nتایی تبدیل می‌شود. از آن‌جا که تعداد حالات قالب kتایی ورودی ۲^k است، تنها ۲^k حالت از ۲^n بردار ممکن در فضای nتایی متعلق به بردارهای کد هستند. در نتیجه هنگامی که تعدادی از بیت‌های یک بردار کد بر اثر نویز جمعی تغییر می‌کنند، بردار نهایی با احتمال زیاد، بردار کد نخواهد بود.

در اکثر کانال‌های مخابراتی، هنگامی که توان سیگنال به اندازه کافی بالا باشد، احتمال تغییر یک بیت توسط نویز، پایین است. از این‌رو در بردار نویزی دریافتی، انتظار تعداد بیت غلط کمی داریم. به عبارت بهتر، اختلاف بین بردارهای ارسالی و دریافتی، یک بردار تنک است. پس، برای بازیابی بردار کد ارسالی، باید نزدیک‌ترین بردار کد به بردار نویزی دریافتی را بیابیم؛ البته در این‌جا، نزدیک و دور بودن بردارها براساس فاصله همینگ[۷] سنجیده می‌شود.

از دیدگاه قابلیت تصحیح خطا، کد مناسب کدی است که فاصله بین بردارهای کد تا حد امکان زیاد باشد؛ معیار سنجش مورد استفاده، اغلب حداقل فاصله بین بردارهای کد است. به عنوان مثال، اگر حداقل فاصله بین بردارهای کد ۵ باشد (حداقل تعداد بیت‌های متفاوت بین دو بردار کد)، به هر نحوی که حداکثر ۴ بیت از یک بردار کد تغییر کنند، بردار نهایی متعلق به فضای کد نخواهد بود و وجود خطا قابل آشکارسازی است. همچنین اگر حداکثر ۲ بیت خطا وجود داشته باشد، بردار کد اولیه در فاصله حداکثر ۲ از بردار نویزی قرار دارد، حال آن‌که سایر بردارهای کد حداقل فاصله ۳ را نسبت به این بردار دارند. بنابراین با یافتن نزدیک‌ترین بردار کد به بردار دریافتی، خطای ایجاد شده به طور صحیح برطرف می‌شود.

نکته جالب توجه در نظریه کدگذاری این است که اگر یک نحوه کدگذاری برای طول ورودی k و طول خروجی n با حداقل فاصله d_{min} بین بردارهای کد وجود داشته باشد، حتماً یک کدگذاری خطی با همین طول‌ها و حداقل فاصله وجود دارد. در نتیجه، از لحاظ قابلیت تصحیح خطا، کدهای خطی بهینه‌اند [۶۶]. در کدگذاری قالبی خطی، ارتباط بین بردار بیت ورودی ($\mathbf{u}_{k\times ۱}$) و بردار بیت خروجی ($\mathbf{c}_{n\times ۱}$) توسط ماتریس مولد کد قابل

[۵] Block Coding
[۶] Parity
[۷] Hamming Distance

بیان است:

$$\mathbf{c}_{n\times 1} = \mathbf{G}_{n\times k} \cdot \mathbf{u}_k \qquad (۲۹-۲)$$

در رابطه ماتریسی فوق، تمام اعمال (جمع و ضرب) در میدان $GF(۲)$ صورت می‌گیرند. همچنین تاثیر نویز بر روی بردار ارسالی را می‌توان توسط بردار خطا ($\mathbf{w}_{n\times 1}$) به صورت زیر بیان کرد:

$$\mathbf{r}_{n\times 1} = \mathbf{c}_{n\times 1} + \mathbf{w}_{n\times 1} \qquad (۳۰-۲)$$

(جمع در میدان $GF(۲)$) از آن‌جا که رتبه ماتریس $\mathbf{G}_{n\times k}$، حداکثر k است، ماتریس $\mathbf{H}_{(n-k)\times n}$ وجود دارد که $\mathbf{H}_{(n-k)\times n} \cdot \mathbf{G}_{n\times k} = \mathbf{0}_{(n-k)\times k}$. این ماتریس به ماتریس آزمون توازن شهرت دارد. اکنون با ضرب کردن طرفین رابطه (۲-۳۰) در $\mathbf{H}_{(n-k)\times n}$ داریم:

$$\mathbf{s}_{(n-k)\times 1} = \mathbf{H}_{(n-k)\times n} \cdot \mathbf{r}_{n\times 1} = \mathbf{H}_{(n-k)\times n} \cdot \mathbf{w}_{n\times 1} \qquad (۳۱-۲)$$

به بیان بهتر، بردار $\mathbf{s}_{(n-k)\times 1}$ تنها متاثر از بردار خطاست. عمل بازیابی بردارکد که هم ارز با یافتن بردار خطاست، در حقیقت معادل به دست‌آوردن بردار تنک $\mathbf{w}_{n\times 1}$ از دستگاه فرو-معین $\mathbf{s}_{(n-k)\times 1} = \mathbf{H}_{(n-k)\times n} \cdot \mathbf{w}_{n\times 1}$ است. این مسأله به وضوح منطبق بر مسأله نمونه‌برداری فشرده است. علاوه بر این، در نظریه کدگذاری، مشابه نمونه‌برداری فشرده، هدف طراحی ماتریس $\mathbf{G}_{n\times n}$ و متعاقباً $\mathbf{H}_{(n-k)\times}$ است به نحوی که قابلیت بازسازی بردارهای تنک خطا را فراهم کند. شایان ذکر است که در هر دو مبحث، ماتریس‌های تصادفی عملکردی مناسب و فراتر از ماتریس‌های یقیینی از خود نشان می‌دهند.

با وجود شباهت بسیار زیاد مسأله نمونه‌برداری فشرده با مسأله طراحی کد، تفاوت عمده در نحوه محاسبات ماتریسی است. در مبحث نمونه‌برداری فشرده، تمامی اعمال در میدان اعداد حقیقی و یا مختلط صورت می‌گیرند در حالی که در نظریه کدگذاری، میدان‌های متناهی مورد توجه هستند. در میدان‌های متناهی، به دلیل ساختار گسسته المان‌ها، مفهوم فاصله قابل تعریف نیست؛ به عبارت دیگر، برخلاف میدان اعداد حقیقی، میدان‌های متناهی توپولوژی مناسبی ندارند. برای روشن‌تر شدن این موضوع، یک دستگاه معادلات خطی n-معادله n-مجهول را در میدان اعداد حقیقی در نظر بگیرید؛ درصورتی که دترمینان ماتریس ضرایب ناصفر باشد، این دستگاه به‌طور یکتا قابل حل است اما بسته به عدد حالت این ماتریس، ممکن است جواب بسیار ناپایدار و یا از نظر عددی، غیرقابل محاسبه باشد. در مقابل، به دلیل متناهی بودن مرتبه میدان، چنین مشکلی

هیچ‌گاه در میدان‌های متناهی رخ نمی‌دهد. برای مثال، ماتریس‌های واندرموند از جمله گزینه‌های خوب برای ماتریس آزمون توازن به‌شمار می‌روند، حال آن‌که به دلیل مشکل ناپایداری، برای ماتریس حسگر انتخاب‌های مناسبی به حساب نمی‌آیند.

۲-۵ مروری بر روش‌های بازسازی

شرط RIP از مرتبه $2k$ با $0 \leq \delta_{2k}$ تضمین می‌کند که در حالت بدون نویز هیچ دو بردار k-تنک‌ای نمونه‌های یکسان تولید نمی‌کنند و در نتیجه بردار k-تنک به طور یکتا از روی نمونه‌ها قابل بازسازی است. اما نحوه‌ی بازسازی بهینه هنوز مشخص نیست. خوشبختانه روش‌هایی که در زیر به آن‌ها اشاره می‌شود، در حد قابل قبولی عمل بازسازی را میسر می‌سازند.

۲-۵-۱ کمینه کردن نرم ℓ_1

همانطور که قبلاً اشاره شد، روش بهینه در بازسازی سیگنال‌های تنک، کمینه کردن نرم ℓ_0 است که در عمل ممکن نیست. به علاوه، اگر چنین کاری ممکن باشد، به دلیل ناپایدار بودن نرم ℓ_0 وجود نویز می‌تواند تفاوت بسزایی در عملکرد این روش ایجاد کند. دو قضیه‌ی زیر نشان می‌دهند که کمینه کردن نرم ℓ_1 می‌تواند تحت شرایطی، جایگزین مناسبی برای این روش باشد.

قضیه ۲-۱ فرض کنید $\mathbf{A}_{m\times n}$ ماتریسی باشد که شرط *RIP* از مرتبه $2k$ را با ثابت $\delta_{2k} < \sqrt{2}-1$ ارضا کند؛ همچنین فرض کنید $\mathbf{x}_{n\times 1}$ بردار دلخواهی باشد و $\mathbf{y}_{m\times 1} = \mathbf{A}_{m\times n}\mathbf{x}_{n\times 1}$. اگر $\mathbf{x}^*_{n\times 1}$ بردار با حداقل نرم ℓ_1 باشد که $\mathbf{y}_{m\times 1} = \mathbf{A}_{m\times n}\mathbf{x}^*_{n\times 1}$ و $\mathbf{x}^{(k)}_{n\times 1}$ بهترین تقریب k-تنک بردار $\mathbf{x}_{n\times 1}$ باشد، داریم:

$$\|\mathbf{x}^* - \mathbf{x}\|_{\ell_2} \leq \frac{C_0}{\sqrt{k}}\|\mathbf{x} - \mathbf{x}^{(k)}\|_{\ell_1} \qquad (۳۲-۲)$$

که C_0 ضریبی ثابت (مستقل از $\mathbf{A}$ و $\mathbf{x}$) است [۲۴].

قضیه فوق نشان می‌دهد که اگر بردار $\mathbf{x}_{n\times 1}$ k-تنک باشد و $\mathbf{A}$ شرط RIP مورد نظر را ارضا کند، کمینه کردن نرم ℓ_1 (BP) این سیگنال را دقیقاً بازسازی خواهد کرد. در [۴۸] نشان داده شده است که نتیجه قضیه فوق برای محدوده وسیع تری از δ_{2k} نیز کماکان معتبر است.

قضیه ۲-۲ فرض کنید $\mathbf{A}_{m\times n}$ ماتریسی باشد که شرط *RIP* از مرتبه $2k$ را با ثابت $0 \le \delta_{2k} < \sqrt{2}-1$ ارضا کند، همچنین فرض کنید $\mathbf{x}_{n\times 1}$ بردار دلخواهی باشد و $\mathbf{y}_{m\times 1} = \mathbf{A}_{m\times n}\mathbf{x}_{m\times 1} + \mathbf{n}_{m\times 1}$ که $\|\mathbf{n}_{m\times 1}\|_{\ell_2} \le \epsilon$ (بردار نویز جمعی). اگر $\mathbf{x}^*_{n\times 1}$ بردار با حداقل نرم ℓ_1 باشد که $\|\mathbf{A}\mathbf{x}^* - \mathbf{y}\|_{\ell_2} \le \epsilon$ و $\mathbf{x}^{(k)}_{n\times 1}$، بهترین تقریب k-تنک بردار $\mathbf{x}_{n\times 1}$ باشد، داریم:

$$\|\mathbf{x}^* - \mathbf{x}\|_{\ell_2} \le \frac{C_0}{\sqrt{k}}\|\mathbf{x} - \mathbf{x}^{(k)}\|_{\ell_1} + C_1\epsilon \qquad (۳۳-۲)$$

که C_0, C_1 ضرایب ثابتی هستند [۲۴].

قضیه فوق نشان می‌دهد که روش کمینه کردن با نرم ℓ_1 نه تنها نسبت به نویز سیستم (k-تنک نبودن بردار) پایدار است، بلکه نسبت به نویز جمعی نیز پایداری خطی دارد. شرط استفاده شده در قضیه ۲-۲ ($\|\mathbf{A}\mathbf{x}^* - \mathbf{y}\|_{\ell_2} \le \epsilon$) به شرط LASSO شهرت دارد [۸۸].

از آن‌جا که نرم ℓ_1 تابعی محدب است، کمینه کردن این نرم را می‌توان توسط روش‌های بهینه‌سازی محدب انجام داد. نکته مهم‌تر در مورد نرم ℓ_1 این است که مسأله کمینه کردن این نرم را می‌توان به صورت یک مسأله خطی بیان کرد. در نتیجه روش‌های برنامه نویسی خطی[۸] به طور موثری در پیدا کردن نقطه بهینه این مسأله می‌توانند مورد استفاده قرار گیرند. روش‌های معروف در این زمینه Simplex و Interior Point هستند که نقطه بهینه را با پیچیدگی محاسباتی از مرتبه n^3 (بعد بردار$\mathbf{x}$) پیدا می‌کنند.

در هر دو مسأله کمینه کردن نرم ℓ_1 (با و بدون نویز) با استفاده از ضرایب لاگرانژ می‌توان نشان داد که بردار حداقل کننده[۹] با بردار حداقل کننده

$$\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2_{\ell_2} + \lambda\|\mathbf{x}\|_{\ell_1} \qquad (۳۴-۲)$$

در صورت انتخاب صحیح λ برابر است. به عنوان مثال، در مسأله نویزی هنگامی‌که واریانس نویز σ_n^2 باشد، $\lambda = \sigma_n\sqrt{2\log n}$ گزینه نزدیکی به مقدار بهینه خواهد بود [۳۱] . با وجود آن که روش برنامه نویس خطی در زمان متناهی جواب مسأله را می‌یابد، در بسیاری از موارد کاربردی، پیچیدگی محاسباتی از مرتبه n^3 غیرعملی است. از این رو، روش‌های دیگری معرفی شده‌اند که به صورت تکراری به سمت جواب بهینه میل می‌کنند. با وجود این که این روش‌ها احتمالاً در زمان متناهی به جواب دقیق دست نمی‌یابند، اما خیلی سریع به جواب

[۸] Linear Programming
[۹] Minimizer

اصلی نزدیک می‌شوند و به کمک این روش‌ها می‌توان در زمانی بسیار کوتاه، به جواب نسبتاً دقیقی دست یافت. از روش‌های معروف در این راستا می‌توان به SPGL1 [۳۷] و GPSR [۴۶] اشاره کرد. در هر تکرار از این روش‌ها، حداقل کننده مسأله زیر محاسبه می‌شود :

$$\arg\min_{\mathbf{x}} \quad \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_{\ell_2} \quad \text{subjectto } \|\mathbf{x}\|_{\ell_1} \leq \tau \tag{۲-۳۵}$$

که τ از روی λ محاسبه می‌شود و در هر تکرار مقدار λ بروز می‌شود.

در هنگامی‌که مقدار λ به نحوی از پیش معلوم باشد، روش‌های بسیار سریعی برای پیدا کردن حداقل کننده (۲-۳۴) وجود دارد. به عنوان مثال، در روش Reweighted LS در هر تکرار و بر حسب تقریب کنونی از جواب، جمله $\|\mathbf{x}\|_{\ell_1}$ با یک جمله مربعی جایگزین می‌شود و در نتیجه در هر تکرار یک مسأله حداقل مربعات حل می‌گردد [۳۵]. روش‌های ISTA و FISTA [۱۲، ۹۴] که اخیراً توجه زیادی را به خود جلب کرده‌اند، بر پایه تعویض جمله $\|\mathbf{A}\mathbf{x}-\mathbf{b}\|_{\ell_2}$ و جایگزاری یک جمله مناسب، مسأله کمینه‌سازی برداری در n بعد را به n مسأله کمینه‌سازی تک بعدی تقلیل می‌دهند. این روش‌ها نیز سرعت بالایی در همگرایی به جواب اصلی دارند .

۲-۵-۲ روش‌های حریص

روش‌های حریص[۱۰] بر پایه این اصل استوار است که در هر مرحله بدون آینده نگری، بهترین انتخاب ممکن صورت گیرد. همان طور که پیش بینی می‌شود، عدم وجود آینده نگری در این روش‌ها، در بسیاری از موارد منجر به شکست می‌شود اما در مورد بازسازی بردارهای تنک موفقیت قابل توجهی کسب کرده‌اند. برای توضیح این دسته از روش‌ها، از ساده‌ترین عضو این خانواده که Matching Pursuit است شروع می‌کنیم. فرض کنید $\mathbf{y}_{m\times 1} = \mathbf{A}_{m\times n}\mathbf{x}_{n\times 1}$ و $\mathbf{x}$ یک بردار k-تنک است. در این صورت $\mathbf{y}_{m\times 1}$ را می‌توان به صورت ترکیب خطی kتا از ستون‌های $\mathbf{A}_{m\times n}$ تلقی کرد و عمل بازسازی معادل با یافتن این k ستون و ضرایب آن‌ها در ترکیب خطی است. بر خلاف روش‌های مبتنی بر نرم ℓ_1 که یک تابع هزینه را کمینه می‌کنند، در این روش ابتدا ستون‌های مورد استفاده از $\mathbf{A}_{m\times n}$ در ترکیب خطی آشکار می‌شوند[۱۱] و سپس به کمک حل یک مسأله حداقل مربعات، ضرایب این ستون‌ها که همان مقادیر ناصفر بردار $\mathbf{x}_{n\times 1}$ هستند محاسبه می‌شوند. در هر مرحله از روش MP یکی از ستون‌های $\mathbf{A}_{m\times n}$ به عنوان عضوی فعالی در ترکیب خطی آشکار می‌شود؛ در این مرحله، ضرب داخلی $\mathbf{y}_{m\times 1}$ با

[۱۰] Greedy

[۱۱] Support Recovery

تمام ستون‌های $\mathbf{A}_{m\times n}$ محاسبه می‌شود و ستونی که بیشترین اندازه را حاصل کند (که بیشترین شباهت را به $\mathbf{y}_{m\times ۱}$ دارد) به عنوان عضو فعال شناسایی می‌شود و مقدار ضرب داخلی که به اندازه این ستون تقسیم شده است را به عنوان ضریب آن لحاظ می‌کنیم. اکنون تقریبی ۱–تنک از بردار اصلی بدست آورده‌ایم که در این مرحله انتخاب بهینه به شمار می‌رود ($\hat{\mathbf{x}}^{(۱)}_{n\times ۱}$) . برای ادامه، با فرض این که محاسبات تا این لحظه صحیح است، بردار $\mathbf{r}_{m\times ۱} = \mathbf{y}_{m\times ۱} - \mathbf{A}_{m\times n}\hat{\mathbf{x}}^{(۱)}_{n\times ۱}$ را به عنوان باقیمانده[۱۲] در نظر می‌گیریم و تمامی‌مراحل قبلی را طی می‌کنیم تا تخمین ۱–تنک جدیدی بدست آید. اکنون حاصل جمع این بردار ۱–تنک جدید و $\hat{\mathbf{x}}^{(۱)}_{n\times ۱}$ را به عنوان تقریب جدید $\hat{\mathbf{x}}^{(۲)}_{n\times ۱}$ تلقی می‌کنیم و این مراحل را مجدداً تا رسیدن به شرط پایانی (نرم باقیمانده کوچک یا تعداد مرحله معلوم و یا ترکیبی از هر دو) ادامه می‌دهیم. این روش از آن‌جا که در به روز رسانی تقریب‌های $\mathbf{x}$ از همبستگی ستون‌های $\mathbf{A}_{m\times n}$ صرف نظر شده است (ستون‌ها عمود فرض می‌شوند) ضعف‌های زیادی دارد. در روش OMP در مرحله‌ی i و پس از آشکار شدن i ستون فعال از ماتریس $\mathbf{A}$ به کمک یک مسأله حداقل مربعات، بهترین ضرایب را برای این بردارها مستقل از نتایج مراحل قبلی حساب می‌کنیم؛ به عبارت بهتر از نتایج قبلی تنها در یافتن مکان‌های ناصفر بردار $\mathbf{x}$ بهره می‌جوییم و نه بیشتر. از آنجا که روش‌های مذکور بر اساس شباهت بردار نمونه‌ها و باقیمانده به ستون‌های $\mathbf{A}$ پایه ریزی شده‌اند، ضریب همدوسی ماتریس $\mathbf{A}$ نقش مهم تری نسبت به پارامتر RIP آن ایفا می‌کند. همچنین روش‌های مذکور در برخی موارد نتایج ضعیفی در حضور نویز از خود نشان می‌دهند. از جمله روش‌های قوی در این خانواده می‌توان از CoSaMP نام برد؛ در این روش هر بار به جای انتخاب یک ستون از ماتریس $\mathbf{A}$ تعداد بیشتری انتخاب می‌شوند که پایداری نسبتاً خوبی را فراهم می‌کند، برای این روش تضمین مشابهی نسبت به کمیته کردن نرم $\ell_۱$ در مورد کیفیت بازسازی هنگام وجود نویز و یا k–تنک نبودن منبع اثبات شده است [۷۴]. روش‌های حریص به طور عام بسیار سریع‌تر از روش‌های دیگر هستند که این سرعت معمولاً به قیمت کیفیت تمام می‌شود.

۲–۵–۳ روش‌های آستانه‌ای

در روش‌های حریص، هر بار شدت ستون‌های ماتریس نسبت به بردار باقیمانده را حساب می‌کنیم و براساس یک قاعده، یک و یا تعدادی از ستون‌ها را به عنوان کاندیداهای ستون‌های فعال شناسایی می‌کنیم. نکته مهم این است که ستون‌های انتخابی در مرحله ثابت است و به روش بستگی دارد . دسته دیگری از روش‌ها وجود دارند

[۱۲] Residual

که مشابه روش‌های حریص به نوعی از شباهت بردارها سود می‌جویند ولی در هنگام انتخاب به جای معیار تعداد از یک سطح آستانه استفاده می‌کنند. یعنی در هر مرحله هر تعدادی از ستون‌ها که شباهت کافی به بردار مورد نظر را دارند انتخاب می‌شوند. در این صورت ممکن است در مرحله‌ای هیچ ستونی انتخاب نشود و یا تمامی‌ستون‌ها انتخاب می‌شوند .

از نماینده‌های معروف این دسته می‌توان به روش IHT[۱۳] اشاره کرد. در این روش برای حل

$$\arg\min_{\mathbf{x}} \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_{\ell_2}^2 + \lambda\|\mathbf{x}\|_{\ell_0} \tag{۳۶-۲}$$

(ضرایب لاگرانژ برای نرم ℓ_0) هر بار از تکرارهای زیر استفاده می‌کنیم:

$$\mathbf{x}^{(i+1)} = \mathcal{H}\big(\mathbf{x}^{(i)} + \mathbf{A}^H(\mathbf{y}-\mathbf{A}\mathbf{x}^{(i)})\big) \tag{۳۷-۲}$$

که منظور از عملگر $\mathcal{H}$ مقایسه با سطح آستانه $\sqrt{\lambda}$ است:

$$\mathcal{H}(x_i) = \begin{cases} 0 & |x_i| \le \sqrt{\lambda} \\ x_i & |x_i| > \sqrt{\lambda} \end{cases} \tag{۳۸-۲}$$

در [۱۵] نشان داده شده است که روش فوق به سمت یکی از حداقل‌های موضعی مسأله (۲-۳۶) میل می‌کند.

در روش IHT هر بار مقایسه با یک سطح آستانه ثابت صورت می‌گیرد. روش IMAT [۷۰، ۸۶] تعمیمی‌از این روش است که سطح آستانه به طور وفقی بر اساس نتایج بدست آمده تا کنون تغییر می‌کند و در نتیجه سرعت همگرایی را افزایش می‌دهد. در [۹۷] نیز روشی مشابه بر مبنای الگوریتم CFAR برای آشکارسازی و حذف نویز ضربه‌ای ارائه شده است.

۲-۵-۴ روش‌های تقریب نرم

دسته کوچکی از روش‌ها مانند SLO[۱۴] [۷۳] براساس تقریب زدن نرم ℓ_0 با توابع دیگر شکل گرفته‌اند. در این روش‌ها در هر تکرار، به جای کمینه کردن نرم ℓ_0، تابع دیگری که احتمالاً یافتن حداقل کننده آن ساده‌تر است مورد استفاده قرار می‌گیرد. در طول تکرارها، توابع مورد استفاده به سمت نرم ℓ_0 نزدیک می‌شوند و در نتیجه جواب نهایی حداقل کننده نرم ℓ_0 خواهد بود. از آن‌جا که نرم ℓ_0 محدب نیست، توابع مورد استفاده نیز نمی‌توانند همگی محدب باشند و در نتیجه حداقل کردن آن‌ها نیز کار دشواری است . نکته اصلی در این روش‌ها این است

[۱۳] Iterative Hard Thresholding

[۱۴] Smoothed ℓ_0

که مقدار حداقل کننده تابع در مرحله کنونی به عنوان نقطه ابتدایی مرحله بعدی مورد استفاده قرار می‌گیرد. از این رو اگر تغییرات توابع به کندی صورت گیرد، هر بار نقطه ابتدایی و نقطه بهینه به اندازه کافی به یکدیگر نزدیک هستند و روش‌های مبتنی بر گرادیان برای یافتن حداقل کننده‌های این توابع غیرمحدب کفایت می‌کنند.

فصل ۳

مروری بر ماتریس‌های حسگر

۳-۱ مقدمه

دو قسمت اصلی در مبحث نمونه‌برداری فشرده عبارتند از عمل نمونه‌برداری و عمل بازسازی. در فصل گذشته به طور مختصر به روش‌های بازسازی اشاره شد. از آن‌جا که نوآوری اصلی این پایان‌نامه در روش‌های نمونه‌برداری است، مرور جامع‌تری در این قسمت ارائه می‌شود. همان طور که در فصل ۱ ذکر شد، عمل نمونه‌برداری در مبحث نمونه‌برداری فشرده به صورت خطی و توسط ماتریس حسگر صورت می‌گیرد. برای بررسی توانایی‌ها و محدودیت‌های ساختاری در این مبحث، به طور تاریخی ابتدا ماتریس‌های تصادفی مورد بررسی قرار گرفته‌اند. به دلیل وجود تقارن در بسیاری از ساختارهای تصادفی، بررسی این ماتریس‌ها ساده‌تر از ماتریس‌های یقینی صورت می‌گیرد. از طرف دیگر، همانند مبحث کدگذاری، برای رسیدن به ظرفیت شبکه، نتایج بدست آمده از ماتریس تصادفی تا کنون قوی‌تر از ماتریس‌های یقینی بوده است. به بیان ساده‌تر، به کمک نتایج حاصل از ماتریس‌های تصادفی، وجود ماتریس‌هایی با پارامترهای مطلوب به اثبات رسیده است در حالی که هنوز پیاده‌سازی برای این گونه ماتریس‌ها یافت نشده است. در این فصل، در دو بخش جداگانه، به بررسی ماتریس‌های تصادفی و یقینی می‌پردازیم. قابل ذکر است که نوآوری اصلی نگارنده در این پایان‌نامه که در فصل ۴ بیان خواهد شد، ارائه ماتریس‌های یقینی است.

۳-۲ ماتریس‌های حسگر تصادفی

برای توضیح در مورد ماتریس‌های تصادفی، از حالت ساده گوسی شروع می‌کنیم. فرض کنید $\mathbf{A}_{m\times n}$ ماتریسی تصادفی با درایه‌های گوسی مستقل از هم‌باشد به نحوی که میانگین هر درایه صفر و واریانس ($\sigma^۲$)آن $\frac{۱}{m}$ باشد. همچنین فرض کنید $\mathbf{x}_{n\times ۱}$ یک بردار یقینی دلخواه باشد و تعریف کنید $\mathbf{y}_{m\times ۱} = \mathbf{A}_{m\times n}\mathbf{x}_{n\times ۱}$. در این صورت داریم:

$$۱ \le i \le m: \quad y_i = \sum_{j=۱}^{n} a_{ij}x_j \tag{۱-۳}$$

از آن‌جا که y_iها ازترکیب خطی متغیرهای گوسی مستقل حاصل شده‌اند، خود گوسی هستند و

$$\begin{cases} \mu_y &= \mathscr{E}\{y_i\} = \sum_{j=۱}^{n} \mathscr{E}\{a_{ij}\}x_j = ۰ \\ \sigma_y^۲ &= \mathscr{E}\{y_i^۲\} = \sum_{j_۱,j_۲=۱} \mathscr{E}\{a_{ij_۱}a_{ij_۲}\}x_{j_۱}x_{j_۲} = \frac{\|\mathbf{x}_{n\times ۱}\|_{\ell_۲}^۲}{m} \end{cases} \tag{۲-۳}$$

همچنین y_iها از یکدیگر مستقل هستند. با توجه به میانگین و واریانس y_iها، $\frac{\sqrt{m}}{\|\mathbf{x}\|_{\ell_۲}}y_i$ یک متغیر گوسی استاندارد است (واریانس واحد) و در نتیجه $l_i = \frac{my_i^۲}{\|\mathbf{x}\|_{\ell_۲}^۲}$ یک متغیر Chi-Square استاندارد با درجه آزادی یک است. داریم:

$$\begin{cases} l &= \sum_{i=۱}^{m} l_i = m\frac{\|\mathbf{y}\|_{\ell_۲}^۲}{\|\mathbf{x}\|_{\ell_۲}^۲} \\ \mu_l &= \mathscr{E}\{l\} = \frac{m^۲}{\|\mathbf{x}\|_{\ell_۲}^۲}\sigma_y^۲ = m \end{cases} \tag{۳-۳}$$

برای بررسی نامساوی‌های مربوط به شرط RIP، احتمال برقراری این نامساوی‌ها را مطالعه می‌کنیم:

$$\begin{cases} P\Big(\frac{\|\mathbf{Ax}\|_{\ell_۲}^۲}{\|\mathbf{x}\|_{\ell_۲}^۲} \le ۱+\delta\Big) &= P\big(l \le m(۱+\delta)\big) = ۱ - P\big(l > m(۱+\delta)\big) \\ P\Big(\frac{\|\mathbf{Ax}\|_{\ell_۲}^۲}{\|\mathbf{x}\|_{\ell_۲}^۲} \ge ۱-\delta\Big) &= P\big(l \ge m(۱-\delta)\big) = ۱ - P\big(l < m(۱-\delta)\big) \end{cases}$$
$$\Rightarrow \quad P\Big(۱-\delta \le \frac{\|\mathbf{Ax}\|_{\ell_۲}^۲}{\|\mathbf{x}\|_{\ell_۲}^۲} \le ۱+\delta\Big) \ge ۱ - P\big(l > m(۱+\delta)\big) - P\big(l < m(۱-\delta)\big) \tag{۴-۳}$$

با استفاده از نامساوی چرنف[۱] برای هر $\nu > ۰$ داریم:

$$\begin{aligned} P\big(l > m(۱+\delta)\big) &< \mathscr{E}_l\Big\{e^{\nu(l-m(۱+\delta))}\Big\} = \mathscr{E}_{\{l_i\}_i}\Big\{e^{\nu\sum_{i=۱}^{m} l_i - (۱+\delta)}\Big\} \\ &= \Big(\mathscr{E}_{l_۱}\Big\{e^{\nu(l-(۱+\delta))}\Big\}\Big)^m = e^{-\nu m(۱+\delta)}\Big(\mathscr{E}_{l_۱}e^{\nu l_۱}\Big)^m \end{aligned} \tag{۵-۳}$$

با توجه به تابع مشخصه توزیع Chi-Square استاندارد با مرتبه واحد، برای $\nu < \frac{۱}{۲}$ داریم:

$$\mathscr{E}_{l_۱}\Big\{e^{\nu l_۱}\Big\} = (۱-۲\nu)^{-\frac{۱}{۲}} = e^{-\frac{۱}{۲}\ln ۱ - ۲\nu} \tag{۶-۳}$$

[۱] Chernoff

که به کمک (۳–۵) نتیجه می‌دهد:

$$P\big(l > m(1+\delta)\big) < e^{-m(\nu(!+\delta)+\frac{1}{2}\ln(1-2\nu))} \tag{۳–۷}$$

با استفاده از مشتق گیری می‌توان نشان داد که بهترین مقدار ν که در نامساوی فوق حداقل کران بالا را ایجاد می‌کند، عبارت است از $\nu = \frac{\delta}{2(1+\delta)}$. پس

$$P\big(l > m(1+\delta)\big) < e^{-\frac{m}{2}(\delta+\ln(1+\delta))} \tag{۳–۸}$$

به طور مشابه می‌توان نشان داد:

$$P\big(l < m(1-\delta)\big) < e^{\frac{m}{2}(\delta+\ln(1-\delta))} \tag{۳–۹}$$

دقت کنید که :

$$\begin{cases} \delta+\ln(1+\delta) & \geq \quad \frac{\delta^2}{2}-\frac{\delta^3}{3} \\ \delta+\ln(1-\delta) & \leq \quad -\frac{\delta^2}{2} \end{cases} \tag{۳–۱۰}$$

بنابراین احتمال برقرار نبودن هر یک از نامساوی‌های RIP مربوط به بردار $\mathbf{x}$ به صورت نمایی نسبت به m کاهش می‌یابد. حال فرض کنید $\mathcal{Q}$ مجموعه‌ای N عضوی از بردارهای n بعدی باشد؛ با استفاده از کران اجتماع خواهیم داشت:

$$\begin{aligned} P\Big(\forall \mathbf{x}\in\mathcal{Q}:\ 1-\delta \leq \frac{\|\mathbf{A}\mathbf{x}\|_{\ell_2}^2}{\|\mathbf{x}\|_{\ell_2}^2} \leq 1+\delta\Big) &\geq 1-N\Big(e^{-\frac{m}{2}(\frac{\delta^2}{2}-\frac{\delta^3}{3})}+e^{-\frac{m\delta^2}{4}}\Big) \\ &\geq 1-2Ne^{-\frac{m}{2}(\frac{\delta^2}{2}-\frac{\delta^3}{3})} \end{aligned} \tag{۳–۱۱}$$

نا منفی بودن کران پایین در نامساوی فوق نشانه آن است که با احتمال ناصفر، ماتریس تصادفی $\mathbf{A}$ وجود دارد که شرط RIP را برای تمام اعضای مجموعه $\mathcal{Q}$ ارضا می‌کند. با استفاده از این روش، می‌توان برقراری نامساوی‌های RIP را به طور همزمان تنها در مورد تعداد متناهی راستا در فضای n بعدی (اگر شرط RIP برای یک بردار $\mathbf{x}_{n\times 1}$ برقرار باشد، برای تمام مضارب آن نیز برقرار است) اثبات کرد. اما تعداد راستاهای بردارهای k–تنک نامحدود است. ایده‌ی جالب توجه در این‌جا این است که می‌توان از برقراری نامساوی‌های RIP برای یک مجموعه‌ی متناهی خاص از بردارهای تنک، برقراری نامساوی برای همه بردارهای تنک را نتیجه گرفت [۱۱].

فرض کنید $\mathcal{T}$ زیرمجموعه‌ای k عضوی از $\{1,2,\ldots,n\}$ باشد و $X_{\mathcal{T}}$ مجموعه تمام بردارهای n بعدی k–تنک باشد که تنها در مکان‌های متعلق به $\mathcal{T}$ عضو ناصفر دارند. ابتدا نامساوی‌های RIP را برای زیرفضای k بعدی

$X_{\mathcal{T}}$ بررسی می‌کنیم. همان طور که پیش‌تر اشاره شد، برقراری و یا عدم برقراری شرط RIP تنها به راستای بردار $\mathbf{x}_{n\times 1}$ بستگی دارد و نه اندازه آن. بنابراین اگر نامساوی‌ها را فقط برای بردارهای با نرم واحد (ابر کره واحد) اثبات کنیم، برای تمامی‌بردارها اثبات کرده‌ایم. با توجه به قضایای مربوط به پوشش کره واحد k- بعدی، می‌دانیم که می‌توان $\lfloor (\frac{۲۴}{\delta})^k \rfloor$ نقطه روی کره انتخاب کرد به نحوی که هر نقطه روی کره به فاصله حداکثر $\frac{\delta}{۸}$ از یکی از این نقاط قرار گیرد [۶۷]. نشان می‌دهیم که اگر نامساوی‌های RIP با ثابت $\frac{\delta}{۴}$ برای این نقاط متعلق به زیر فضای $X_{\mathcal{T}}$ برقرار باشند، نامساوی‌های RIP با ثابت δ برای تمام زیر فضای $X_{\mathcal{T}}$ برقرارند. فرض کنید $\mathbf{x}_{n\times 1} \in X_{\mathcal{T}}$ برداری با نرم واحد باشد که بزرگترین ثابت δ_u را در نامساوی سمت راست شرط RIP برای تمام بردارهای $X_{\mathcal{T}}$ دارد (حالت $\delta_u \geq ۱$ نیز در این‌جا قابل قبول است). همچنین فرض کنید $\mathbf{q}_{n\times 1}$ نزدیک‌ترین نقطه در بین مجموعه مذکور به $\mathbf{x}_{n\times 1}$ باشد:

$$\begin{aligned}\sqrt{۱+\delta_u} = \|\mathbf{A}_{m\times n}\mathbf{x}_{n\times 1}\|_{\ell_۲} &\leq \|\mathbf{A}_{m\times n}\mathbf{q}_{n\times 1}\|_{\ell_۲} + \|\mathbf{A}_{m\times n}(\mathbf{x}_{n\times 1}-\mathbf{q}_{n\times 1})\|_{\ell_۲} \leq \sqrt{۱+\frac{\delta}{۲}} + \sqrt{۱+\delta_u}\frac{\delta}{۸} \\ &\Rightarrow \quad ۱+\delta_u \leq \frac{۱+\frac{\delta}{۲}}{(۱-\frac{\delta}{۸})^۲} \quad \Rightarrow \quad \delta_u \leq \frac{\frac{۳}{۴}\delta - \frac{۱}{۶۴}\delta^۲}{(۱-\frac{۱}{۸})^۲} < \delta \end{aligned} \tag{۱۲-۳}$$

حال فرض کنید $\mathbf{x}_{n\times 1} \in X_{\mathcal{T}}$ یک بردار دلخواه باشد و $\mathbf{q}_{n\times 1}$ نزدیک نقطه در بین مجموعه مورد نظر به این نقطه باشد، داریم:

$$\begin{aligned}\|\mathbf{A}_{m\times n}\mathbf{x}_{n\times 1}\|_{\ell_۲} &\geq \|\mathbf{A}_{m\times n}\mathbf{q}_{n\times 1}\|_{\ell_۲} - \|\mathbf{A}_{m\times n}(\mathbf{x}_{n\times 1}-\mathbf{x}_{n\times 1})\|_{\ell_۲} \geq \sqrt{۱-\frac{\delta}{۲}} - \sqrt{۱+\delta}\frac{\delta}{۸} \\ &\geq ۱-\frac{\delta}{۲} - (۱+\delta)\frac{\delta}{۸} \geq ۱-\delta \end{aligned} \tag{۱۳-۳}$$

پس شرط RIP با ثابت δ برای تمام زیر فضای $X_{\mathcal{T}}$ برقرار است. در روند بالا، ابتدا یک زیرمجموعه k عضوی $\mathcal{T}$ از $\{۱,\dots,n\}$ انتخاب کردیم و سپس برای ایجاد شرط RIP روی کل زیر فضای $X_{\mathcal{T}}$ تعداد $(\frac{۲۴}{\delta})^k$ نقطه را برای برقراری شرط RIP با ثابت $\frac{\delta}{۴}$ انتخاب کردیم. اگر بخواهیم حکم فوق را به کل زیرمجموعه‌های k عضوی از $\{۱,\dots,n\}$ تعمیم دهیم، کافی است شرط RIP با ثابت $\frac{\delta}{۴}$ برای مجموعه‌ای $(\frac{۲۴}{\delta})^k\binom{n}{k}$ عضوی برقرار شود. با توجه به (۳-۱۱) احتمال برقراری شرط RIP با ثابت $\frac{\delta}{۴}$ برای این مجموعه حداقل برابر است با:

$$۱ - ۲\Big(\frac{۲۴}{\delta}\Big)^k \binom{n}{k} e^{-m(\frac{\delta^۲}{۱۶}-\frac{\delta^۳}{۲۴})} \tag{۱۴-۳}$$

با استفاده از تقریب استرلینگ و در نظر گرفتن $k \ll n$ داریم:

$$\binom{n}{k} = \mathcal{O}\Big((\frac{n}{k})^k (\frac{n}{n-k})^{n-k}\sqrt{\frac{n}{k(n-k)}}\Big) \leq e^{\mathcal{O}(k\ln\frac{n}{k}+k)} \tag{۱۵-۳}$$

در نتیجه احتمال برقراری شرط RIP با ثابت $\frac{\delta}{۲}$ روی مجموعه مفروض به طور حدی از

$$۱ - ۲e^{\mathcal{O}\left(k(\ln\frac{n}{k}+۱+\ln\frac{۲۴}{\delta})-m(\frac{\delta^۲}{۱۶}-\frac{\delta^۳}{۲۴})\right)} \qquad (۱۶-۳)$$

کمتر نیست. پس اگر $m > \mathcal{O}(k\ln\frac{n}{k})$، توان عدد منفی و بزرگی خواهد بود و در نتیجه کران پایین روی احتمال بسیار به یک نزدیک خواهد بود. به بیان ساده‌تر اگر $m > \mathcal{O}(k\ln\frac{n}{k})$، احتمال برقراری شرط RIP مرتبه‌ی k (با هر ثابت δ) بسیار نزدیک به یک خواهد بود. این نتیجه نه تنها وجود ماتریس‌های یقینی که شرط RIP را ارضا می‌کنند نشان می‌دهد بلکه بیان می‌کند که هر تحقق از ماتریس تصادفی گوسی با احتمال نزدیک به یک این شرط را برآورده خواهد کرد.

با کمی‌دقت در روابط بدست آمده می‌توان به این نتیجه رسید که تنها استفاده‌ای که از گوسی‌بودن توزیع احتمال شده، این است که اگر $\mathbf{A}_{m\times n}$ با درایه‌های i.i.d. و توزیع گوسی اختیار شوند:

$$P\left(\left|\|\mathbf{A}_{m\times n}\mathbf{x}_{n\times ۱}\|_{\ell_۲}^۲ - \mathscr{E}\{\|\mathbf{A}_{m\times n}\mathbf{x}_{n\times ۱}\|_{\ell_۲}^۲\}\right| \geq \epsilon\mathscr{SD}\{\|\mathbf{A}_{m\times n}\mathbf{x}_{n\times ۱}\|_{\ell_۲}^۲\}\right) \leq ۲e^{-mf(\epsilon)} \qquad (۱۷-۳)$$

که f یک تابع مثبت و کراندار و $\mathscr{SD}\{.\}$ معرف انحراف معیار است (جذر واریانس). نامساوی فوق به نامساوی تمرکز اندازه[۲] شهرت دارد و تنها محدود به توزیع گوسی نیست. حالت کلی‌تری از آن چه در بالا برای توزیع گوسی ذکر شد به نام قضیه Johnson-Lindenstrauss تعمیم‌یافته به صورت زیر وجود دارد [۱، ۵۶، ۵۹]:

قضیه ۳-۱ فرض کنید $۰ < \delta < ۱$ یک ثابت داده شده و $\mathcal{Q}$ یک مجموعه N عضوی از $\mathbb{R}^n$ باشد. در این صورت برای هر $m > m_۰ = \mathcal{O}(\frac{\ln N}{\delta^۲})$ تابع خطی $f : \mathbb{R}^n \to \mathbb{R}^m$ وجود دارد که برای هر $\mathbf{u}_{n\times ۱}, \mathbf{v}_{n\times ۱} \in \mathcal{Q}$:

$$(۱-\delta)\|\mathbf{u}-\mathbf{v}\|_{\ell_۲}^۲ \leq \|f(\mathbf{u})-f(\mathbf{v})\|_{\ell_۲}^۲ \leq (۱+\delta)\|\mathbf{u}-\mathbf{v}\|_{\ell_۲}^۲ \qquad (۱۸-۳)$$

در حقیقت شرط RIP حالت خاصی از قضیه فوق است که $\mathcal{Q}$ شامل بردار تمام صفر است و ما تنها به دنبال نامساوی‌هایی هستیم که یکی از بردارهای $\mathbf{u}$ و $\mathbf{v}$ همین بردار تمام صفر است.

۳-۳ ماتریس‌های حسگر غیرتصادفی

در بخش گذشته با استفاده از ماتریس‌های تصادفی نشان دادیم که اگر تعداد سطرهای ماتریس (m) بیشتر از یک کران پایین با مرتبه $k\ln\frac{n}{k}$ باشد، شرط RIP قابل دستیابی است. در این بخش به دنبال روش‌های ساخت ماتریس

[۲] Measure Concentration

حسگر هستیم به نحوی که بتوان وجود شرط RIP را تضمین کرد. برای این منظور مجدداً شرط RIP را بررسی می‌کنیم.

اگر ماتریس $\mathbf{A}_{m\times n}$ شرط RIP مرتبه k را با ثابت $0 \leq \delta_k < 1$ ارضا کند، برای تمامی‌بردارهای k–تنک مثل $\mathbf{x}_{n\times 1}$ می‌دانیم که $\|\mathbf{Ax}\| > 0$. در نتیجه هیچ بردار k–تنکی در فضای پوچ ماتریس قرار نمی‌گیرد. در ماتریس‌های واندرموند با k سطر می‌دانیم که هر k انتخاب از ستون‌های ماتریس مستقل خطی اند. یعنی اگر $\mathbf{A}$ یک ماتریس واندرموند باشد، برای هر بردار k–تنک مثل $\mathbf{x}$ خواهیم داشت $\|\mathbf{Ax}\| > 0$. اما آیا این به منزله‌ی برقراری شرط RIP است؟ برای بررسی بیشتر فرض کنید $\{\mathbf{a}_i\}_{i=1}^n$ ستون‌های ماتریس $\mathbf{A}$ و $\{x_i\}_{i=1}^n$ درایه‌های بردار $\mathbf{x}$ باشند به نحوی که تنها $\{x_{i_1}, \ldots, x_{i_k}\}$ ناصفر اند. در این صورت:

$$\mathbf{A}_{m\times n}\mathbf{x}_{n\times 1} = \sum_{j=1}^{k} x_{i_j}\mathbf{a}_{i_j} \tag{۱۹–۳}$$

اگر تعریف کنیم $\mathbf{A}_{m\times k}^{(\text{sub})} = [\mathbf{a}_{i_1} \ \ldots \ \mathbf{a}_{i_k}]$، شرط RIP بر روی ماتریس $\mathbf{A}$ ایجاب می‌کند که:

$$\forall \mathbf{x}_{k\times 1} \in \mathbb{R}^k : \quad 1 - \delta_k \leq \frac{\|\mathbf{A}_{m\times k}^{(\text{sub})}\mathbf{x}_{k\times 1}\|_{\ell_2}^2}{\|\mathbf{x}_{k\times 1}\|_{\ell_2}^2} \leq 1 + \delta_k \tag{۲۰–۳}$$

با کمک نتایج جبرخطی می‌دانیم که عبارت $\frac{\|\mathbf{A}_{m\times k}^{(\text{sub})}\mathbf{x}_{k\times 1}\|_{\ell_2}^2}{\|\mathbf{x}_{k\times 1}\|_{\ell_2}^2}$ همواره بین مربعات حداقل و حداکثر مقادیر تکین[۳] تغییر می‌کند و دو حالت حدی را نیز اختیار می‌کند. به بیان ساده‌تر، شرط RIP برای ماتریس $\mathbf{A}$ محدوده تغییرات مقادیر تکین تمام زیر ماتریس‌های k ستونی آن را تعیین می‌کند. به طور خاص، عدد حالت[۴] تمام زیر ماتریس‌های k ستونی $\mathbf{A}$ باید از کرانی که تنها به δ_k وابسته است کمتر باشند. متاسفانه عدد حالت ماتریس‌های واندرموند به صورت نمایی با افزایش m زیاد می‌شوند [۴۹] و در نتیجه ماتریس واندرموند شرط RIP را تنها برای مقادیر δ_k نزدیک به یک می‌تواند ارضا کند (برای mهای بزرگ). هنگامی‌که δ_k افزایش یابد، فاصله‌ی نسبی بردارهای تکین با فضای پوچ ماتریس کاهش می‌یابند و در نتیجه عمل بازسازی نمونه‌ها از جایی به بعد ناپایدار خواهد شد. به این دلیل گزینه امیداور کننده‌ی ماتریس‌های واندرموند را باید کنار گذاشت.

از اولین روش‌های ساخت یقینی ماتریس حسگر می‌توان به [۵۴] اشاره کرد. در این مقاله به کمک کدهای مرتبه دوم Reed-Muller یک ماتریس حسگر $2^l \times 2^{\frac{l(l+1)}{2}}$ با درایه‌های ± 1 معرفی شده است که ساختاری

[۳] Singular Value

[۴] Condition Number

به صورت

$$\mathbf{A}_{\mathrm{RM}} = \left[\mathbf{U}_1\ \mathbf{U}_2\ \dots\ \mathbf{U}_{2^{\frac{l(l-1)}{2}}}\right] \tag{۲۱-۳}$$

دارد که هر کدام از $\mathbf{U}_i$ها یک ماتریس مربعی $2^l \times 2^l$ با ستون‌های متعامد است. در واقع برای زیر ماتریس‌های $\mathbf{U}_i$ از ماتریس $\mathbf{A}_{\mathrm{RM}}$ تمامی‌مقادیر تکین یکسانند که باعث ایجاد حالت ایده آل $\delta_{2^l} = 0$ در این زیر ماتریس‌ها می‌شود. متاسفانه در مورد بقیه‌ی زیر ماتریس‌های $\mathbf{A}_{\mathrm{RM}}$ نتیجه‌ی قابل توجه‌ای وجود ندارد و در نتیجه نمی‌توان برقراری شرط RIP را در این ماتریس تضمین کرد؛ با این حال نتایج شبیه‌سازی‌ها نشان می‌دهد که این ماتریس عملکرد مناسبی دارد.

یکی از مثال‌های معروف از ماتریس‌های یقینی، ماتریس‌های بر مبنای توابع Chirp هستند [۹]. فرض کنید m یک عدد طبیعی است؛ هر عدد بین ۱ تا m^2 را می‌توان به صورت $\alpha m + \beta$ که $0 \le \alpha \le m-1$ و $1 \le \beta \le m$ نمایش داد. ماتریس $m \times m^2$ بر مبنای توابع Chirp ($\mathbf{A}_{\mathrm{chirp}}$) با درایه‌های a_{ik} به صورت زیر ساخته می‌شود:

$$\begin{cases} 1 \le i \le m \\ 1 \le k = \alpha m + \beta \le m^2 \end{cases} \Rightarrow\ a_{ik} = e^{j\frac{2\pi}{m}(\alpha i^2 + \beta i)} \tag{۲۲-۳}$$

ماتریس حاصل درایه‌های مختلط با اندازه واحد (قبل از نرمالیزه کردن با $\frac{1}{\sqrt{m}}$) و سطرهای عمود بر هم‌دارد:

$$\langle \mathrm{row}_{i_1}, \mathrm{row}_{i_2} \rangle = \sum_{\alpha=0}^{m-1} \sum_{\beta=1}^{m} e^{j\frac{2\pi}{m}(\alpha i_1^2 + \beta i_1 - \alpha i_2^2 - \beta i_2)} = \sum_{\alpha=0}^{m-1} e^{j\frac{2\pi}{m}\alpha(i_1^2 - i_2^2)} \underbrace{\sum_{\beta=1}^{m} e^{j\frac{2\pi}{m}\beta(i_1 - i_2)}}_{0} = 0 \tag{۲۳-۳}$$

در [۹] نشان داده شده است که این ماتریس تا حدی خواص ماتریس‌های واندرموند را دارد، به این صورت که در ماتریس $m \times m^2$، هر $\sqrt{m}$ ستون مستقل خطی اند و در نتیجه اگر این ماتریس برای نمونه‌برداری از یک بردار k–تنک با $k \le \frac{\sqrt{m}+1}{2}$ مورد استفاده قرارگیرد، نمونه‌های بدست آمده به طور یکتا بردار تنک را مشخص می‌کنند. متاسفانه هیچ‌گونه تضمینی برای برقراری شرط RIP در این ماتریس‌ها به اثبات نرسیده است. از نکات مثبت این ماتریس‌ها روش بازسازی بسیار ساده‌ی آنهاست؛ برای بازسازی سیگنال تنک می‌توان به نوعی از تبدیل DFT بردار خود همبستگی نمونه‌ها استفاده کرد و مکان‌های ناصفر سیگنال تنک را آشکار کرد. به دلیل استفاده از الگوریتم FFT پیچیدگی محاسباتی در این روش بسیار پایین است.

در فصل ۴ نشان خواهیم داد که اگر ضریب همدوسی ماتریس $\mathbf{A}$ کمتر از $\frac{1}{k-1}$ باشد، این ماتریس شرط RIP را برآورده خواهد کرد. تاکنون این مطلب تنها روش تضمین شرط RIP در ماتریس‌ها بوده است. همان

طور که در فصل ۱ اشاره شد، با استفاده از این مطلب نمی‌توان به ماتریس‌هایی با ابعادی که ماتریس‌های تصادفی پیشنهاد می‌کنند رسید. از مثال‌های ساده ماتریس با ضریب همبستگی کم می‌توان به‌ترکیب ماتریس همانی و ماتریس DFT اشاره کرد. فرض کنید $\mathbf{I}_{m\times m}$ و $\mathbf{F}_{m\times m}$ به‌ترتیب ماتریس همانی و ماتریس یکانی تبدیل فوریه باشند و قرار دهید $\mathbf{A}_{m\times ۲m}=\left[\mathbf{I}_{m\times m}\ \mathbf{F}_{m\times m}\right]$. فرض کنید $\mathbf{a}_{m\times ۱},\mathbf{b}_{m\times ۱}$ دو ستون متمایز از این ماتریس باشند. اگر هر دو بردار متعلق به یکی از ماتریس‌های $\mathbf{I}_{m\times m}$ یا $\mathbf{F}_{m\times m}$ باشند، بر هم‌عمود خواهند بود. در غیر این صورت یکی از دو بردار (مثلاً $\mathbf{a}_{m\times ۱}$) متعلق به $\mathbf{I}_{m\times m}$ و دیگری متعلق به $\mathbf{F}_{m\times m}$ است. از آن‌جا که $\mathbf{a}_{m\times ۱}$ تنها یک درایه ناصفر دارد و اندازه تمامی‌المان‌های $\mathbf{b}_{m\times ۱}$ برابر $\frac{۱}{\sqrt{m}}$ است، داریم:

$$\left|\langle \mathbf{a}_{m\times ۱},\mathbf{b}_{m\times ۱}\rangle\right|=\frac{۱}{\sqrt{m}} \tag{۲۴-۳}$$

در نتیجه علاوه بر این که ستون‌های $\mathbf{A}$ یکه‌اند، اندازه ضرب داخلی هر جفت از آن‌ها حداکثر $\frac{۱}{\sqrt{m}}$ است، پس $\mu_{\mathbf{A}}=\frac{۱}{\sqrt{m}}$. این بدان معنی است که ماتریس $\mathbf{A}$ شرط RIP از مرتبه $k=\lceil\sqrt{m}\ \rceil$ را برآورده می‌کند. در مثال فوق، ماتریس‌های همانی و DFT دو ماتریس یکانی بودند که اندازه ضرب داخلی هر دو ستون از آن‌ها مقدار ثابت $\frac{۱}{\sqrt{m}}$ بود. به چنین ماتریس‌هایی غیرهمدوس می‌گویند؛ برای واضح‌تر شدن این موضوع، فرض کنید $\mathbf{C}_{m\times m}$ و $\mathbf{D}_{m\times m}$ دو ماتریس یکانی باشند و $\{\mathbf{c}_i\}_{i=۱}^{m}$ و $\{\mathbf{d}_i\}_{i=۱}^{m}$ به‌ترتیب ستون‌های $\mathbf{C}$ و $\mathbf{D}$ باشند. از آن‌جا که $\mathbf{C}$ یکانی است، در رابطه پارسوال[۵] صدق می‌کند:

$$۱=\|\mathbf{d}_i\|_{\ell_۲}^{۲}=\|\mathbf{C}\mathbf{d}_i\|_{\ell_۲}^{۲}=\sum_{j=۱}^{m}|\langle \mathbf{c}_j,\mathbf{d}_i\rangle|^{۲} \tag{۲۵-۳}$$

در نتیجه حداقل برای یکی از $\mathbf{c}_j$ها، $|\langle \mathbf{c}_j,\mathbf{d}_i\rangle|\geq\frac{۱}{\sqrt{m}}$. پس $\frac{۱}{\sqrt{m}}$ یک کران پایین برای حداکثر اندازه ضرب داخلی بین ستون‌های دو ماتریس یکانی است. به وضوح تساوی هنگامی رخ می‌دهد که اندازه ضرب داخلی هر ستون از $\mathbf{C}_{m\times m}$ با هر ستون از $\mathbf{D}_{m\times m}$ مقداری ثابت باشد؛ از این رو، در این حالت به دو ماتریس غیرهمدوس می‌گویند. زوج‌های غیرهمدوس تنها محدود به ماتریس همانی و DFT نمی‌شوند اما برای تمام این حالات، ماتریس ادغامی‌حاصل $m\times ۲m$ خواهد بود.

تعمیم‌هایی از روش فوق به بیش از $۲m$ ستون نیز مطرح شده است؛ به عنوان مثال، در حالاتی خاص با کنار هم‌قرار دادن چند ماتریس یکانی می‌توان به ماتریس با ضریب همدوسی کوچک دست یافت، اما از آن‌جا که به دنبال کاهش ضریب همدوسی ماتریس هستیم به جای کنار هم قرار دادن ماتریس‌های یکانی بهتر است

[۵] Parseval

ستون‌ها را به نحو مناسب‌تری انتخاب کنیم. یکی از ابزارهای قوی در این راستا فریم‌های گراسمانی[۶] هستند [۸۷]. به مجموعه $\{\mathbf{a}_i\} \subset \mathbb{R}^m$ فریم گوییم هرگاه ضرایب $0 < \alpha \leq \beta < \infty$ وجود داشته باشند به طوری که برای هر $\mathbf{x} \in \mathbb{R}^m$،

$$\alpha\|\mathbf{x}\|_{\ell_2}^2 \leq \sum_i |\langle \mathbf{a}_i, \mathbf{x}\rangle|^2 \leq \beta\|\mathbf{x}\|_{\ell_2}^2 \qquad (۲۶-۳)$$

در واقع یک فریم مجموعه‌ای از بردارهاست که توسط ضرب داخلی، نرم بردارها را تا حدی ثابت نگه می‌دارند. در صورتی که $\alpha = \beta$، فریم اصطلاحاً تنگ[۷] نامیده می‌شود. در این فریم‌ها به نوعی قضیه پارسوال صادق است؛ در نتیجه از این جهت بسیار شبیه پایه‌های متعامد یکه هستند. نوع دیگری از فریم‌ها وجود دارند که از جهت ضریب همدوسی به پایه‌های متعامد یکه شبیه‌اند: در صورتی که به بردارهای یکه در فضای m بعدی محدود باشیم و بخواهیم m بردار به نحوی انتخاب کنیم که ضریب همدوسی بین آن‌ها حداقل شود، پایه متعامد یکه با ضریب همدوسی صفر جواب خواهد بود. حال اگر بخواهیم بیش از m بردار انتخاب کنیم، حتماً ضریب همدوسی بزرگتر از صفر خواهد بود. به مجموعه n عضوی از بردارهای m بعدی $(m \leq n)$ با نرم واحد که کوچکترین ضریب همدوسی ممکن بین تمامی‌حالات را دارند، فریم گراسمانی گویند. در فصل ۱ بیان کردیم که $\sqrt{\frac{n-m}{m(n-1)}}$ یک کران پایین برای ضریب همدوسی این مجموعه است (نامساوی ولش) اما آیا می‌توان به این کران دست یافت؟ نکته جالب در این‌جاست که اگر ضریب همدوسی مجموعه‌ای برابر با این کران پایین باشد، حتماً فریم تنگ است [۸۷]. علاوه بر این، این کران تنها برای $n \leq \frac{m(m+1)}{2}$ قابل دستیابی است و وجود چنین مجموعه‌هایی در این حالت نیز اثبات شده است اما به غیر از حالات $\frac{n}{m} \leq 2$، ساخت این مجموعه‌ها مستلزم جستجو است. یعنی نه تنها فرم بسته‌ای برای این ساختارها در دسترس نیست، بلکه تعداد اعضا از تابعی درجه ۲ نسبت به بعد فضا کوچکتر است. این نتیجه اختلاف عمیقی بین ابعاد پیشنهادی ماتریس‌های تصادفی (رابطه لگاریتمی n با m) و این ساختارها را روشن می‌کند.

از جمله موفق‌ترین طرح‌ها در ماتریس‌های حسگر، روش ارائه شده در [۳۸] است. در این روش به کمک چند جمله ای‌ها در میدان متناهی، ماتریس‌های $p^2 \times p^{r+1}$ طراحی شده‌اند که شرط RIP از مرتبه $k < \frac{p}{r} + 1$ را ارضا می‌کند. دقت کنید که در این روش، تعداد ستون‌ها $(n = p^{r+1})$ لزوماً به تابعی درجه ۲ از تعداد سطرها $(m = p^2)$ محدود نیست؛ البته این ساختار به قیمت فاصله گرفتن از کران ولش حاصل شده است. در زیر به

[۶] Grassmanian Frames

[۷] Tight

طور مختصر نحوه‌ی ساخت ماتریس بیان می‌شود:

فرض کنید p توانی از یک عدد اول و $GF(p)$ میدان متناهی با مرتبه p باشد. فرض کنید $Q(x)$ یک چندجمله‌ای با ضرایب $GF(p)$ باشد و $\mathcal{G}(Q)$ مجموعه تمام زوج‌های مرتب به صورت $\big(x, Q(x)\big)$ که $x \in GF(p)$؛ یعنی $\mathcal{G}(Q)$ یک مجموعه p عضوی است. حال فرض کنید $\mathcal{B}$ مجموعه تمام زوج‌های مرتب $GF(p) \times GF(p)$ باشد که به وضوح $p^۲$ عضوی خواهد بود. برای سادگی نمایش، فرض کنید $\mathcal{B} = \{b_۱, \dots, b_{p^۲}\}$ که هر b_i یک زوج مرتب است و زوج‌ها باترتیب واژه ای[۸] مرتب شده‌اند. بـرای هـر چنـد جمله‌ای مثل Q، یک بردار $p^۲$ تایی $(\mathbf{v}_Q)$ با درایه‌های ۰،۱ به صورت زیر تعریف می‌شود:

$$\mathbf{v}_Q = [v_۱, \dots, v_{p^۲}]^T \quad , \quad v_i = \begin{cases} ۰ & b_i \notin \mathcal{G}(Q) \\ ۱ & b_i \in \mathcal{G}(Q) \end{cases} \tag{۳–۲۷}$$

بنابراین $\mathbf{v}_Q$ دقیقاً شامل p تا یک و $p(p-۱)$ صفر است. اکنون فرض کنید $\mathcal{P}_r$ مجموعه تمام چند جمله‌ای‌های $GF(p)[x]$ با درجه حداکثر r باشد:

$$\mathcal{P}_r = \big\{ a_۰ + a_۱ x + \dots + a_r x^r \;\big|\; a_i \in GF(p) \big\} \tag{۳–۲۸}$$

در این صورت $\mathcal{P}_r$ یک مجموعه $p^{r+۱}$ عضوی خواهد بود:

$$\mathcal{P}_r = \{ Q_۱, Q_۲, \dots, Q_{p^{r+۱}} \} \tag{۳–۲۹}$$

از آنجا که درجه چند جمله ای‌ها حداکثر r است، تساوی $Q_i(x) = Q_j(x)$ حداکثر برای r مقدار x اتفاق می‌افتد و در نتیجه $\langle \mathbf{v}_{Q_i}, \mathbf{v}_{Q_j} \rangle \le r$. پس اگر بردارهای $\mathbf{v}_{Q_i}$ را با ضریب $\frac{۱}{\sqrt{p}}$ یکه کنیم مجموعه‌ای از بردارهای یکه با ضریب همدوسی $\frac{r}{p}$ بدست می‌آوریم که برای تشکیل ماتریسی با شرط RIP از مرتبه $k < \frac{p}{r} + ۱$ کفایت می‌کنند.

در روش‌های مبتنی بر ضریب همدوسی همواره داریم $m \ge \mathcal{O}(k^۲)$. در چند ساختار ارائه شد به کمک گراف، ماتریس‌هایی معرفی شده‌اند که $m = \mathcal{O}(k)$، ولی به جای شرط RIP شرط‌های ضعیف‌تری وجود دارند که کماکان عمل بازسازی را میسر می‌سازند. به عنوان مثال، در [۵۵] به کمک ماتریس مجاورت گراف‌های Extractor ماتریس‌های $m = k۲^{\mathcal{O}(\log\log n)^۲}$ ارائه شده‌اند که قادر به بازیابی سیگنال‌های k–تنک هستند. گراف دو بخشی G را با بخش‌های A و B به طوری که $|A| = n$ و $|B| = k$ در نظر بگیرید. فرض کنید درجه تمام رأس‌های A برابر با $d_A = ۲^{\mathcal{O}(\log\log n)^۲}$ باشد و برای هر زیرمجموعه از رأس‌های G مانند X فرض کنید $\Gamma(X)$

[۸] Lexicographic

همسایه‌های X در G باشند. این گراف را ϵ-Extractor گوییم هرگاه برای هر زیرمجموعه حداقل k عضوی از A داشته باشیم:

$$\sum_{b\in B}\left|\frac{|\Gamma(b)\cap X|}{d_A\cdot|X|}-\frac{1}{k}\right|\leq\epsilon \qquad (۳۰-۳)$$

شرط فوق به طور ضمنی بیان می‌کند که یالهای مربوط به رأس‌های مجموعه‌های k عضوی (و بیشتر) از A درجه تقریباً یکسانی در رأس‌های B ایجاد می‌کنند. نحوه ساخت چنین ماتریس‌های برای $\epsilon > 0$ قبلاً بررسی شده است. در مقاله [۵۵] بر اساس ماتریس مجاورت گراف‌های $\frac{1}{16}$-Extractor، ماتریس‌های حسگری معرفی شده است که بدون برقراری شرط RIP قابلیت بازسازی سیگنالهای k–تنک را فراهم می‌کنند؛ البته نحوه بازسازی ارائه شده تنها هنگامی‌کارآمد است که نمونه‌ها بدون نویز باشند.

براساس گراف‌های Expander طرح دیگری برای ماتریس‌های حسگر در [۵۷] با ابعاد $m=\mathcal{O}(k\log\frac{n}{k})$ معرفی شده‌اند. با وجود این که این طرح به صورت ساختاری ارائه شده است، در ساختن ماتریس نیاز به گراف‌هایی است که خود با جستجو بدست می‌آیند. نکته قوت این روش در نحوه تایید ماتریس است؛ در حالت کلی بررسی شرط RIP در ماتریس مسأله NP به شمار می‌رود، حال آن که در طرح ارائه شده بررسی مناسب بودن گراف در زمان چندجمله‌ای صورت می‌پذیرد. همچنین ماتریس مجاورت این گراف‌ها که به عنوان ماتریس حسگر مورد استفاده قرار می‌گیرند، به جای نرم ℓ_2، شرط RIP با نرم ℓ_1 را ارضا می‌کنند. بررسی‌های انجام شده در [۵۷] نشان می‌دهند که این شرط نیز می‌تواند بازسازی پایدار را تضمین کند. نکته قابل توجه در این طرح روش بازسازی است؛ به دلیل استفاده از گراف‌های Expander می‌توان بازسازی را به کمک الگوریتم Belief Propagation انجام داد.

همان گونه که در ابتدا ذکر شد، شرط RIP از مرتبه k با ثابت δ_k ایجاب می‌کند که مربع مقادیر تکین تمام زیر ماتریس‌های k ستونی ماتریس اصلی در بازه $[1-\delta_k, 1+\delta_k]$ قرار گیرند. در صورتی که حتی فقط برای یک زیر ماتریس این خاصیت برقرار نباشد، ماتریس شرط RIP را ندارد، اما ممکن است قادر به بازسازی بسیاری از بردارهای k–تنک باشد. برای بیان قابلیت ناقص بازسازی، شرط RIP احتمالاتی در [۱۹] مطرح شده است. گوییم ماتریس $\mathbf{A}_{m\times n}$ شرط StRIP با پارامترهای (k,δ,ϵ) را ارضا می‌کند هرگاه نامساوی‌های

$$(1-\delta)\|\mathbf{x}\|_{\ell_2}^2\leq\left\|\frac{1}{\sqrt{m}}\mathbf{A}_{m\times n}\mathbf{x}_{n\times 1}\right\|_{\ell_2}^2\leq(1+\delta)\|\mathbf{x}\|_{\ell_2}^2 \qquad (۳۱-۳)$$

با احتمال $1-\epsilon$ برای بردارهای k–تنک برقرار باشند که در آن بردارهای k–تنک و هم‌نرم، هم‌احتمال فرض شده‌اند. بوضوح شرط StRIP ضعیف‌تر از RIP است و در نتیجه دسته بزرگتری از ماتریس‌ها را شامل می‌شود. یکی از نتایج مهم خاصیت StRIP این است که شرط کافی ساده‌تری نسبت به RIP برای آن وجود دارد. فرض کنید سه شرط زیر در مورد ماتریس $\mathbf{A}$ صادق باشند:

۱. سطرهای $\mathbf{A}$ بر یکدیگر عمودند و جمع درایه‌های هر سطر صفر است؛

۲. ستون‌های $\mathbf{A}$ تحت عمل ضرب درایه به درایه تشکیل گروه می‌دهند، به طور خاص ستون اول $\mathbf{A}$ تمام یک است؛

۳. جمع اندازه درایه‌های هر ستون به غیر از ستون تمام یک، حداکثر $m^{1-\frac{\eta}{2}}$ است که $\eta > 0/5$؛

در این صورت ثابت c وجود دارد که اگر $k < 1 + (n-1)\eta$ و $m \geq \left(c\frac{k \log n}{\delta^2}\right)^{\frac{1}{\eta}}$، ماتریس $\mathbf{A}_{m\times n}$ شرط StRIP با پارامترهای (k, δ, ϵ) را ارضا خواهد کرد که

$$\epsilon = 2e^{-\frac{(\delta - \frac{k-1}{n-1})^2 n\eta}{8k}} \qquad (۳۲-۳)$$

برخلاف شرط RIP، بررسی سه خاصیت ذکر شده به سادگی میسر است و در نتیجه اثبات StRIP بسیار ساده‌تر از RIP صورت می‌گیرد. از جمله ماتریس‌هایی که شرط StRIP را ارضا می‌کنند ماتریس‌های بر مبنای توابع Chirp هستند.

فصل ۴

ماتریس‌های حسگر غیرتصادفی پیشنهادی

۴-۱ مقدمه

همان طور که در فصل گذشته مطرح شد، برای پیاده‌سازی عمل نمونه‌برداری فشرده، نیازمند ماتریس‌های تعیین هستیم . از آن‌جا که شرط RIP بازسازی کامل و پایدار را ایجاب می‌کند، طراحی ماتریس‌های تعیینی که شرط RIP را ارضا می‌کنند بسیار مطلوب است. همان طور که در ادامه شرح داده می‌شود، بر آورده کردن شرط RIP در ماتریس‌های تعیینی، تاکنون تنها توسط استدلالات مبتنی بر ضریب همدوسی ماتریس میسر شده است. نکته منفی در این استدلال، کاهش بعد ماتریس به بهای ارضای شرطی قوی‌تر از RIP (ضریب همدوسی پایین) است. در این فصل به طراحی ماتریس‌های تعیینی با ضریب همدوسی پایین می‌پردازیم. برای این منظور، ابتدا به کمک کدهای متعامد نوری[۱] ماتریس‌های دودویی[۲] طراحی می‌کنیم و به کمک کران جانسون[۳] محدودیت ذاتی ابعاد این ماتریس‌ها را نشان می‌دهیم. سپس به کمک کدهای دودویی خطی، به ویژه کدهای BCH ماتریس‌های دوقطبی(± 1)[۴] و در نتیجه ماتریس‌های مختلط ارائه می‌دهیم. برای دستیابی به ابعاد بزرگتر و تنوع بیشتر، روش‌هایی برای ترکیب ماتریس‌های معرفی شده ارائه می‌دهیم. به طور خاص نشان می‌دهیم که با ترکیب ماتریس‌های دودویی و دوقطبی می‌توان به ماتریس‌هایی با ابعادی فراتر از طرح‌های پیشین دست یافت. در انتها، نشان می‌دهیم که با استفاده از روش‌های بازسازی حریص[۵] می‌توان یک بردار تنک را از روی نمونه‌های بدست

---

[۱] Optical Orthogonal Codes

[۲] Binary

[۳] Johnson's Bound

[۴] Bipolar

[۵] Greedy

آمده توسط این ماتریس‌ها و با پیچیدگی محاسباتی کمتر از میزان رایج $\mathcal{O}(n^۲)$ بدست آورد. دلیل اصلی کاهش محاسباتی، خاصیت چرخشی ستون‌های این ماتریس است که به کمک الگوریتم FFT منجر به تسریع محاسبات ضرب ماتریسی می‌شود. نتایج شبیه‌سازی‌ها نشان می‌دهند که عملکرد ماتریس‌های معرفی شده با عملکرد ماتریس‌های تصادفی قابل مقایسه است با این تفاوت که الگوریتم کارایی برای بازسازی بردار تنک از روی نمونه‌های تصادفی وجود ندارد.

۴-۲ ماتریس‌های دودویی

همان طور که اشاره شد، روش ما در این فصل برای برقرار کردن شرط RIP در یک ماتریس، کاهش ضریب همدوسی آن است. لم زیر، رابطه‌ی بین ضریب همدوسی و شرط RIP را نشان می‌دهد.

لم ۴-۱ اگر ستون‌های ماتریس $\mathbf{A}_{m\times n}$ همگی یکه باشند و ضریب همدوسی ماتریس برابر با $\mu_{\mathbf{A}}$ باشد، در این صورت ماتریس $\mathbf{A}$ ، شرط RIP از مرتبه $k < \frac{۱}{\mu_{\mathbf{A}}} + ۱$ با ثابت $\delta_k = (k-۱)\mu_{\mathbf{A}}$ را ارضا می کند.

اثبات: فرض کنید $\mathbf{x}_{n\times ۱}$ یک بردار با حداکثر k درایه ناصفر (k- تنک) باشد. در این صورت:

$$\|\mathbf{A}\mathbf{x}\|_۲^۲ \quad = \quad \|\sum_{i=۱}^{n} x_i\mathbf{a}_i\|_۲^۲ = \sum_{i,j=۱}^{n} x_i x_j^* \langle \mathbf{a}_i, \mathbf{a}_j\rangle = \|\mathbf{x}\|_۲^۲ + \sum_{i\neq j} x_i x_j^* \langle \mathbf{a}_i, \mathbf{a}_j\rangle \qquad (۱-۴)$$

که در آن $\mathbf{a}_i$ و x_i به ترتیب بیانگر ستون i ماتریس $\mathbf{A}$ و المان i بردار $\mathbf{x}$ هستند، با استفاده از تعریف ضریب همدوسی ماتریس داریم:

$$|\sum_{i\neq j} x_i x_j^* \langle \mathbf{a}_i, \mathbf{a}_j\rangle| \quad \leq \quad \mu_{\mathbf{A}} \sum_{i\neq j} |x_i|\cdot|x_j| = \mu_{\mathbf{A}} \Big(\sum_{i=۱}^{n} |x_i|\Big)^۲ - \mu_{\mathbf{A}}\|\mathbf{x}\|_۲^۲ \qquad (۲-۴)$$

از آن‌جا که بردار $\mathbf{x}$ ، k-تنک است می‌توانیم از نامساوی $\Big(\sum_{i=۱}^{n} |x_i|\Big)^۲ \leq k\|\mathbf{x}\|_۲^۲$ استفاده کنیم تا رابطه زیر بدست آید:

$$|\sum_{i\neq j} x_i x_j^* \langle \mathbf{a}_i, \mathbf{a}_j\rangle| \leq (k-۱)\mu_{\mathbf{A}}\|\mathbf{x}\|_۲^۲ \qquad (۳-۴)$$

که شرط RIP را نتیجه می‌دهد:

$$۱ - (k-۱)\mu_{\mathbf{A}} \leq \frac{\|\mathbf{A}\mathbf{x}\|_۲^۲}{\|\mathbf{x}\|_۲^۲} \leq ۱ + (k-۱)\mu_{\mathbf{A}} \qquad (۴-۴)$$

و در نتیجه اثبات کامل است. ■

ماتریس‌های حسگر دودویی، ماتریس‌هایی هستند که قبل از نرمالیزه کردن ستون‌ها، از درایه‌های ۰ و ۱ تشکیل شده‌اند. یک نحوه ساخت چنین ماتریس‌هایی در [۳۸] معرفی شده است که پیش از این در فصل ۳ بیان شد. در این‌جا، ارتباط بین ماتریس‌های حسگر و کدهای متعامد نوری را بررسی می‌کنیم. کدهای متعامد نوری مجموعه‌ای از بردارهای دودویی با وزن ثابت (تعداد ۱ ثابت) هستند که ضرب داخلی هر دو بردار (و حتی بردارهای حاصل از گردش دوری آن‌ها) نسبت به وزن بردارها مقدار کمی است. از آنجا که درایه‌های این بردارها تنها شامل ۰ و ۱ است، جملات موجود در بسط ضرب دو بردار، همواره نامنفی هستند و در نتیجه با افزایش وزن بردارها، انتظار داریم که ضرب داخلی آن‌ها نیز افزایش یابد. فرض کنید $R(m,w,\lambda)$ حداکثر تعداد بردارهای دودویی $m\times 1$ با وزن w و ضرب داخلی حداکثر برابر با $\lambda\in\mathbb{Z}$ باشد (در این‌جا تنها ضرب داخلی همان بردارها منظور است و نه انتقال[۶] یافته‌های آن‌ها). در [۵۸] یک کران بالای قوی برای $R(m,w,\lambda)$ به صورت زیر اثبات شده است:

$$R(m,w,\lambda)\leq\left\lfloor\frac{m}{w}\left\lfloor\frac{m-1}{w-1}\lfloor\dots\lfloor\frac{m-\lambda}{w-\lambda}\rfloor\dots\rfloor\right\rfloor\right\rfloor \qquad (۵-۴)$$

که در آن $\lfloor x\rfloor$ جزء صحیح x را نشان می‌دهد.

در مخابرات نوری این بردارهای کد به عنوان امضاهای[۷] کاربرهای مختلف تلقی می‌شوند و از آن‌جا که در گیرنده از همبستگی[۸] سیگنال دریافتی با امضای هر کاربر پیام ارسالی آشکار می‌شود، ضرب داخلی کم بین امضاهای مختلف اهمیت بسزایی دارد. از این‌رو، با وجود عمود نبودن بردارهای امضا، به آن‌ها کدهای متعامد نوری (OOC[۹]) می‌گویند [۸۱]. از آن‌جا که ارتباطات نوری اغلب به صورت غیرهمزمان[۱۰] صورت می‌گیرد، نه تنها ضرب داخلی بین امضاها بلکه ضرب داخلی بین هر امضا و گردش‌های دوری سایر امضاها و همچنین گردش‌های دوری خود این امضا اهمیت شایانی دارد. بنابراین به جای یک λ ساده، دو پارامتر λ_a و λ_c تعریف می‌شوند: λ_a برابر است با حداکثر میزان خودهمبستگی در بین بردارهای کد، هنگامی که میزان گردش دوری ناصفر باشد و λ_c حداکثر همبستگی بین هر دو حالت چرخش یافته بردارهای کد مختلف را نشان می‌دهد. کدهای متعامد نوری اغلب به صورت $(m,w,\lambda_a,\lambda_c)$ شناسایی می‌شوند اما هنگامی که دو پارامتر λ_a و λ_c برابر

[۶] shift
[۷] Signature
[۸] Correlation
[۹] Optical Orthogonal Code
[۱۰] Asynchronous

باشند، از نمایش (m, w, λ) استفاده می‌شود.

فرض کنید $\mathcal{A}$ مجموعه‌ای از بردارهای متعامد نوری با بعد m ، وزن w و پارامترهای $\lambda_a = \lambda_c = \lambda$ باشد؛ همچنین فرض کنید تمام گردش‌های دوری این بردارها نیز در $\mathcal{A}$ گنجانده شده باشند. طبق تعریف کدهای متعامد نوری، ضرب داخلی هر دو بردار داخل $\mathcal{A}$ حداکثر λ است. اکنون اگر ماتریس $\mathbf{A}_{m\times n}$ را با قراردادن این بردارها به عنوان ستون‌های ماتریس و سپس یکه کردن آن‌ها با تقسیم ستون‌ها بر $\sqrt{w}$ بسازیم، ضریب همدوسی ماتریس $\mathbf{A}$ حداکثر $\frac{\lambda}{w}$ خواهد بود. در اینجا فرض کرده‌ایم که تعداد اعضاء مجموعه $\mathcal{A}$ برابر n است. همچنین ترتیب قرار گرفتن این بردارها به عنوان ستون‌های $\mathbf{A}$ اهمیتی در ضریب همدوسی ماتریس حاصل ندارد. در زیر به طور مختصر نحوه‌ی ساخت یک کد متعامد نوری را که بر گرفته از [۳۹] و بر مبنای میدان‌های متناهی است، بیان می‌کنیم.

فرض کنید $q = ۱۶^a$ که $a \in \mathbb{N}$ است و همچنین فرض کنید $\mathbb{F} = GF(q)$ و α یک ریشه اولیه میدان $\mathbb{F}$ باشد. از آن‌جا که $۵|q-۱$، می‌توان نوشت $q = ۵d+۱$. تعریف می‌کنیم :

$$\mathcal{D}_i = \{\alpha^{d+i}, \alpha^{۲d+i}, \ldots, \alpha^{۵d+i}\} \quad , \quad ۰ \le i \le d-۱ \tag{۴-۶}$$

از آن‌جا که تعداد ۱ها در کدهای متعامد نوری بسیار کمتر از تعداد ۰هاست، معمولاً بردارهای کد را توسط مکان ۱های آن‌ها مشخص می‌کنیم. در طرح مورد نظر، d بردار کد با طول $q-۱$ می‌سازیم به گونه‌ای که اگر مکان‌های ناصفر بردار iام مجموعه $\mathcal{C}_i$ را تشکیل دهد، داشته باشیم:

$$\mathcal{C}_i = \log_\alpha(\mathcal{D}_i - ۱) \quad , \quad ۱ \le i \le d-۱ \tag{۴-۷}$$

دقت کنید که $۱ \in \mathcal{D}_۰$ و $\mathcal{D}_۰$ برای ساختن کد استفاده نشده است. در [۳۹] نشان داده شده است که با این روش $\frac{۱۶^a-۶}{۵}$ بردار کد تولید می‌شوند که توسط سه تایی $(۱۶^a - ۱, ۵, ۲)$ مشخص می‌شوند. از آن‌جا که در ساختار ماتریس حسگر تمام گردش‌های این بردارها نیز مورد استفاده قرار می‌گیرند، ماتریسی با ابعاد $n \times (۱۶^a - ۱)$ ساخته می‌شود که $n \lessapprox \frac{(۱۶^a-۱)(۱۶^a-۶)}{۵}$. با توجه به ضریب همدوسی، این ماتریس، RIP مرتبه ۳ را با ثابت $\delta_۳ = ۰/۸$ ارضا می‌کند. در [۳۹] با استفاده از روشی مشابه، کدهای متعامد نوری $(۲, w, n)$ با wهای بزرگتر (مرتبه RIP بزرگتر) معرفی شده‌اند که با توجه به نامساوی (۴-۵) شبه بهینه به نظر می‌رسند.

همان طور که پیشتر نیز اشاره شد، در [۳۸] یک طرح ساخت ماتریس دودویی با ابعاد $p^۲ \times p^{r+۱}$ و ضریب همدوسی $\frac{r}{p}$ بدون استفاده از کدهای نوری معرفی شده است. در اینجا p توان صحیحی از یک عدد اول است و همان طور که در فصل ۳ توضیح داده شد، ستون‌های ماتریس بر اساس چندجمله‌ای‌ها در میدان متناهی $GF(p)$ ساخته می‌شود. در زیر نشان می‌دهیم که این ماتریس‌ها، در بین ماتریس‌های دودویی، از لحاظ ضریب همدوسی در حالت حدی $\frac{p}{r^۲} \to \infty$ بهینه هستند:

$$\begin{aligned} \lim_{\frac{p}{r^۲}\to\infty} \frac{p^{r+۱}}{R(p^۲,p,r)} &\geq \lim_{\frac{p}{r^۲}\to\infty} \prod_{i=۰}^{r} \frac{p(p-i)}{p^۲-i} \geq \lim_{\frac{p}{r^۲}\to\infty} \left(\frac{p(p-r)}{p^۲-r}\right)^{r+۱} \geq \lim_{\frac{p}{r^۲}\to\infty} \left(۱-\frac{r}{p}\right)^{r+۱} \\ &= \lim_{\frac{p}{r^۲}\to\infty} \left(\left(۱-\frac{r}{p}\right)^{-\frac{p}{r}}\right)^{-\frac{r(r+۱)}{p}} \geq \lim_{\frac{p}{r^۲}\to\infty} e^{-\frac{r(r+۱)}{p}} = e^{۰} = ۱ \end{aligned} \qquad (۸-۴)$$

۴-۳ ماتریس‌های دو قطبی به کمک کدهای BCH

در این بخش به بررسی ارتباط میان نظریه‌ی کدگذاری و ماتریس‌های حسگر می‌پردازیم. از آن‌جا که پارامترهای k و n هم در مبحث نمونه‌برداری فشرده و هم در نظریه‌ی کدگذاری مورد استفاده قرار می‌گیرند، برای ایجاد تمایز، متغیرهای مربوط به کد را با علامت $\tilde{}$ نمایش می‌دهیم؛ به عنوان مثال $\tilde{n}$ طول کد را نشان میدهد در حالی که n بیانگر تعداد ستونهای ماتریس حسگر است.

فرض کنید $\mathcal{C}(\tilde{n},\tilde{k};۲)$ یک کد خطی بلوکی دودویی و $۱_{\tilde{n}\times ۱}$ بردار تمام یک باشند؛ می‌گوییم $\mathcal{C}$ متقارن است هر گاه $۱_{\tilde{n}\times ۱} \in \mathcal{C}$ باشد. برای کدهای متقارن اگر $\mathbf{a}_{\tilde{n}\times ۱}$ یک بردار کد باشد، به دلیل خطی بودن کد، مکمل آن که به صورت $\mathbf{a}_{\tilde{n}\times ۱} \oplus ۱_{\tilde{n}\times ۱}$ تعریف می‌شود نیز یک بردار کد است. در نتیجه مجموعه بردارهای کد به صورت زوج‌های مکمل خواهند بود.

قضیه ۴-۱ فرض کنید $\mathcal{C}(\tilde{n},\tilde{k};۲)$ یک کد متقارن با حداقل فاصله‌ی $\tilde{d}_{min}$ باشد و فرض کنید $\tilde{\mathbf{A}}_{\tilde{n}\times ۲^{\tilde{k}-۱}}$ ماتریس حاصل از کنار هم گذاشتن بردارهای کد به صورت ستونی باشد به نحوی که از هر زوج مکمل، دقیقاً یکی انتخاب شده باشد. تعریف می کنیم:

$$\mathbf{A}_{\tilde{n}\times ۲^{\tilde{k}-۱}} \triangleq \frac{۱}{\sqrt{\tilde{n}}}\left(۲\tilde{\mathbf{A}}_{\tilde{n}\times ۲^{\tilde{k}-۱}} - (۱)_{\tilde{n}\times ۲^{\tilde{k}-۱}}\right) \qquad (۹-۴)$$

در این صورت، ضریب همدوسی ماتریس $\mathbf{A}$ حداکثر برابر با $\frac{\tilde{n}-۲\tilde{d}_{min}}{\tilde{n}}$ خواهد بود.

اثبات: ابتدا دقت کنید که ستونهای ماتریس $\mathbf{A}$ همگی یکه هستند. در حقیقت، $۲\tilde{\mathbf{A}}_{\tilde{n}\times ۲^{\tilde{k}-۱}} - (۱)_{\tilde{n}\times ۲^{\tilde{k}-۱}}$

همان ماتریس $\mathbf{A}$ است که در آن صفرها با ۱− جایگذاری شده‌اند (نمایش دو قطبی). بنابراین، قدر مطلق هر درایه ماتریس $\mathbf{A}$ برابر با $\frac{۱}{\sqrt{\tilde{n}}}$ است که یکه بودن ستون‌ها را نشان می‌دهد.

فرض کنید $\mathbf{a}_{\tilde{n}\times ۱}, \mathbf{b}_{\tilde{n}\times ۱}$ دو ستون متمایز در ماتریس $\mathbf{A}$ متناظر با ستون‌های $\tilde{\mathbf{a}}_{\tilde{n}\times ۱}, \tilde{\mathbf{b}}_{\tilde{n}\times ۱}$ در ماتریس $\tilde{\mathbf{A}}$ باشند. اگر $\tilde{\mathbf{a}}$ و $\tilde{\mathbf{b}}$ در l مکان متفاوت باشند، داریم:

$$\langle \mathbf{a}, \mathbf{b} \rangle = \frac{۱}{\tilde{n}} \Big(۱ \times (\tilde{n} - l) + (-۱) \times l \Big) = \frac{\tilde{n} - ۲l}{\tilde{n}} \qquad (۱۰\text{–}۴)$$

همچنین $\tilde{\mathbf{b}}$ و $\tilde{\mathbf{a}} \oplus ۱_{\tilde{n}\times ۱}$ (مکمل $\tilde{\mathbf{a}}$) در $\tilde{n} - l$ مکان با یکدیگر متفاوت هستند و از آن‌جا که هر سه بردار $\{\tilde{\mathbf{a}},\ \tilde{\mathbf{a}} \oplus ۱_{\tilde{n}\times ۱},\ \tilde{\mathbf{b}}\}$، بردار کدهای متفاوتی هستند (از هر زوج مکمل دقیقاً یکی انتخاب شده بود) هر دوی l و $\tilde{n} - l$ کمتر یا مساوی $\tilde{d}_{min}$ خواهند بود. یعنی:

$$\begin{cases} l \geq \tilde{d}_{min} \\ \tilde{n} - l \geq \tilde{d}_{min} \end{cases} \quad \Rightarrow \quad \tilde{d}_{min} \leq l \leq \tilde{n} - \tilde{d}_{min} \Rightarrow |\tilde{n} - ۲l| \leq \tilde{n} - ۲\tilde{d}_{min} \qquad (۱۱\text{–}۴)$$

دقت کنید $۰_{\tilde{n}\times ۱}, ۱_{\tilde{n}\times ۱} \in \mathcal{C}$ و برای هر بردار کد $\tilde{\mathbf{a}}$ یکی از دو مقدار $d(۰_{\tilde{n}\times ۱}, \tilde{\mathbf{a}})$ و یا $d(۱_{\tilde{n}\times ۱}, \tilde{\mathbf{a}})$ حداکثر برابر است با $\frac{\tilde{n}}{۲}$. از این رو، $\tilde{n} - ۲\tilde{d}_{min} \geq ۰$. با ترکیب کردن (۴–۱۰) و (۴–۱۱) داریم:

$$|\langle \mathbf{a}, \mathbf{b} \rangle| \leq \frac{\tilde{n} - ۲\tilde{d}_{min}}{\tilde{n}} \qquad (۱۲\text{–}۴)$$

که کران مورد نظر برای ضریب همدوسی ماتریس $\mathbf{A}$ را نتیجه می‌دهد. ■

قضیه فوق تنها زمانی کارساز است که $\tilde{d}_{min}$ نزدیک به $\frac{\tilde{n}}{۲}$ باشد؛ برای این که کران بالای k که از ضریب همدوسی نتیجه می‌شود به اندازه کافی بزرگ باشد، باید صورت کسر ذکر شده در (۴–۱۲) به صفر نزدیک باشد. با توجه به قضایای موجود در نظریه‌ی کدگذاری می‌دانیم که $\tilde{d}_{min} \leq \tilde{n} - \tilde{k} + ۱$. لذا برای این که $\tilde{d}_{min}$ در حد $\frac{\tilde{n}}{۲}$ باشد، باید تعداد بیت‌های توازن حداقل در حد تعداد بیت‌های اطلاعاتی باشند که در مورد کدهای رایج در مخابرات چنین امری دور از ذهن است. در ادامه نشان می‌دهیم که می‌توان به کمک کدهای BCH چنین کدهایی را طراحی کرد.

۴–۳–۱ کدهای BCH با $\tilde{d}_{min}$ بزرگ

کدهای BCH زیردسته‌ای از کدهای چرخشی دودویی هستند که در آن‌ها $\tilde{n} = ۲^{\tilde{m}} - ۱$ $(\tilde{m} \in \mathbb{N})$ و بردارهای کد توسط یک چندجمله‌ای مولد $g(x) \in GF(۲)[x]$ با شرط $g(x) | x^{۲^{\tilde{m}}-۱} + ۱$ تولید می‌شوند [۶۶] . با توجه به

ساختار میدان‌های متناهی داریم:

$$x^{۲^{\tilde{m}}-۱}+۱=\prod_{\substack{r\in GF(۲^{\tilde{m}})\\ r\neq ۰}}(x-r) \qquad (۱۳-۴)$$

در نتیجه چندجمله‌ای مولد در BCH را می‌توان به فاکتورهای خطی در $GF(۲^{\tilde{m}})[x]$ تجزیه کرد. فرض کنید $\alpha\in GF(۲^{\tilde{m}})$ یک ریشه اولیه برای این میدان و α^i یکی از از ریشه‌های $g(x)$ باشد؛ از آن‌جا که $g(x)\in GF(۲)[x]$، تمام مزدوج‌های عنصر α^i نسبت به $GF(۲)$ نیز ریشه‌های $g(x)$ هستند. در میدان‌های متناهی با مشخصه ۲ می‌دانیم که این مزدوج‌ها عناصر متفاوت مجموعه $\{\alpha^{i۲^j}\}_{j=۰}^{\tilde{m}-۱}$ هستند. به علاوه، از آن‌جا که $\alpha^{۲^{\tilde{m}}-۱}=۱$ برای هر $i_۱,i_۲$ که $i_۱\equiv i_۲(\bmod ۲^{\tilde{m}}-۱)$ ، داریم $\alpha^{i_۱}=\alpha^{i_۲}$ که خاصیت چرخشی توان‌ها را نشان می‌دهد.

مزیت اصلی کدهای BCH نسبت به سایر کدهای چرخشی، وجود کران پایین تضمین شده برای $\tilde{d}_{min}$ است: اگر $\alpha^{i_۱},\ldots,\alpha^{i_d}$ ریشه‌های متمایزی از $g(x)$ باشند (نه لزوماً تمامی ریشه‌ها) به طوری که $i_۱,\ldots,i_d$ یک تصاعد حسابی تشکیل دهند، داریم $\tilde{d}_{min}\geq d+۱$.

اکنون به مسئله طراحی کد با $\tilde{d}_{min}$ بزرگ بازمی‌گردیم. در این روش، به جای طراحی چندجمله‌ای مولد، چندجمله‌ای آزمون توازن[۱۱] را می‌سازیم:

$$h(x)\triangleq\frac{x^{۲^{\tilde{m}}-۱}+۱}{g(x)} \qquad (۱۴-۴)$$

به بیان بهتر، هر عنصری از میدان یا ریشه‌ای از $g(x)$ است و یا $h(x)$. برای ساختن $h(x)$ ریشه‌های آن را معرفی می‌کنیم. فرض کنید $l<\tilde{m}-۱$ عددی صحیح است و :

$$\mathcal{G}_{\tilde{m}}^{(l)}=\{\alpha^{۰},\alpha^{۱},\ldots,\alpha^{۲^{\tilde{m}-۱}+۲^l-۱}\} \qquad (۱۵-۴)$$

دقت کنید که تعریف $\mathcal{G}_{\tilde{m}}^{(l)}$ به انتخاب ریشه اولیه α بستگی دارد. اکنون $\mathcal{H}_{\tilde{m}}^{(l)}$ را به صورت زیرمجموعه‌ای از $\mathcal{G}_{\tilde{m}}^{(l)}$ تعریف می‌کنیم که نسبت به عمل مزدوج‌گیری بسته باشد:

$$\mathcal{H}_{\tilde{m}}^{(l)}\triangleq\{r\in\mathcal{G}_{\tilde{m}}^{(l)}\mid\forall\, j\in\mathbb{N}:\ r^{۲^j}\in\mathcal{G}_{\tilde{m}}^{(l)}\} \qquad (۱۶-۴)$$

تعریف فوق نشان می‌دهد که برای هر $r\in\mathcal{H}_{\tilde{m}}^{(l)}$ تمام مزدوج‌های آن به صورت $r^{۲^j}$ نیز داخل $\mathcal{H}_{\tilde{m}}^{(l)}$ قرار دارند.

[۱۱] Parity Check

اکنون چندجمله‌ای آزمون توازن را به صورت زیر تعریف می‌کنیم:

$$h(x) = \prod_{r \in \mathcal{H}_{\tilde{m}}^{(l)}} (x - r) \tag{۱۷-۴}$$

همانطور که اشاره شد، برای هر r که ریشه $h(x)$ باشد، تمام مزدوج‌های آن نیز ریشه $h(x)$ خواهند بود. در نتیجه $h(x) \in GF(۲)[x]$ که شرط لازم و کافی برای دودویی بودن ضرایب $g(x)$ است. همچنین:

$$۱ = \alpha^{۰} \in \mathcal{G}_{\tilde{m}}^{(l)} \Rightarrow ۱ \in \mathcal{H}_{\tilde{m}}^{(l)} \Rightarrow (۱ + x) \big| h(x) \tag{۱۸-۴}$$

که نشان می‌دهد بردار تمام یک متعلق به فضای کد است:

$$\begin{aligned} c = [\underbrace{۱, \ldots, ۱}_{۲^{\tilde{m}}-۱}]^T \quad &\Rightarrow \quad c(x) = ۱ + x + \cdots + x^{۲^{\tilde{m}}-۲} = \frac{x^{۲^{\tilde{m}}-۱} + ۱}{x + ۱} \\ &\Rightarrow \quad x^{۲^{\tilde{m}}-۱} + ۱ \big| (x^{۲^{\tilde{m}}-۱} + ۱) \frac{h(x)}{۱ + x} = c(x) h(x) \end{aligned} \tag{۱۹-۴}$$

در نتیجه کد تولید شده توسط $g(x) = \frac{x^{\tilde{n}}+۱}{h(x)}$ یک کد متقارن است که تمام شرایط قضیه ۴-۱ را بر آورده می‌کند. برای یافتن حداقل فاصله کد، دقت کنید که ریشه‌های $h(x)$ زیرمجموعه‌ای از $\mathcal{G}_{\tilde{m}}^{(l)}$ هستند و در نتیجه، تمام اعضای $GF(۲^{\tilde{m}}) \backslash \mathcal{G}_{\tilde{m}}^{(l)}$ ریشه‌های $g(x)$ هستند:

$$\forall \;\; ۲^{\tilde{m}-۱} + ۲^{l} \leq j \leq ۲^{\tilde{m}} - ۲ : \qquad g(\alpha^{j}) = ۰ \tag{۲۰-۴}$$

بنابراین، یک تصاعد حسابی با طول حداقل $۲^{\tilde{m}-۱} - ۲^{l} - ۱$ در بین توان‌های αای که ریشه $g(x)$ هستند یافت می‌شود. بنابراین:

$$\tilde{d}_{min} \geq (۲^{\tilde{m}-۱} - ۲^{l} - ۱) + ۱ = ۲^{\tilde{m}-۱} - ۲^{l} \tag{۲۱-۴}$$

روند رایج در نظریه‌ی کدگذاری چنین است که بر حسب $\tilde{n}, \tilde{k}$ داده شده، به دنبال کدی با حداکثر $\tilde{d}_{min}$ هستیم . در این‌جا، برای یک $\tilde{d}_{min}$ مناسب و یک $\tilde{n}$ داده شده، کدی طراحی کردیم که $\tilde{k}$ آن همچنان نامعلوم است:

$$\begin{aligned} \tilde{n} &= \tilde{k} + deg\big(g(x)\big) \\ \Rightarrow \tilde{k} &= \tilde{n} - deg\big(g(x)\big) = \big(deg\big(g(x)\big) + deg\big(h(x)\big)\big) - deg\big(g(x)\big) = deg\big(h(x)\big) = |\mathcal{H}_{\tilde{m}}^{(l)}| \end{aligned} \tag{۲۲-۴}$$

قضیه زیر تعداد اعضای مجموعه $|\mathcal{H}_{\tilde{m}}^{(l)}|$ را به یک مساله ترکیبیاتی مربوط می‌کند:

قضیه ۴-۲ با استفاده از نمادهای قبلی، $|\mathcal{H}_{\tilde{m}}^{(l)}|$ برابر است با تعداد دنباله‌های دودویی به طول $\tilde{m}$ به طوری که بین هر دو ۱ متوالی، حداقل $\tilde{m}-l-1$ صفر به صورت گردشی وجود داشته باشد.

اثبات: نشان می‌دهیم که یک نگاشت یک به یک بین تعداد اعضای $\mathcal{H}_{\tilde{m}}^{(l)}$ و دنباله‌های دودویی مذکور وجود دارد. فرض کنید $(b_{\tilde{m}-1},\ldots,b_0)\in\{0,1\}^{\tilde{m}}$ یکی از این دنباله‌های دودویی و β نمایش اعشاری این دنباله باشد:

$$\beta=(\overline{b_{\tilde{m}-1}\ldots b_0})_2=\sum_{i=0}^{\tilde{m}-1}b_i 2^i \qquad (۴-۲۳)$$

نشان می‌دهیم $\alpha^\beta\in\mathcal{H}_{\tilde{m}}^{(l)}$. برای سادگی فرض کنید β_j نمایش اعشاری دنباله اصلی (که β را تولید می‌کند) پس از j واحد انتقال چرخشی به سمت چپ باشد:

$$\begin{aligned}
\beta_0 &= (\overline{b_{\tilde{m}-1}\ldots b_0})_2\\
\beta_1 &= (\overline{b_{\tilde{m}-2}\ldots b_0 b_{\tilde{m}-1}})_2\\
\beta_2 &= (\overline{b_{\tilde{m}-3}\ldots b_0 b_{\tilde{m}-1} b_{\tilde{m}-2}})_2\\
&\vdots\\
\beta_{\tilde{m}-1} &= (\overline{b_0 b_{\tilde{m}-1}\ldots b_1})_2
\end{aligned} \qquad (۴-۲۴)$$

اکنون داریم

$$\begin{aligned}
2\beta_j &= 2\times(\overline{b_{\tilde{m}-1-j}\ldots b_0 b_{\tilde{m}-1} b_{\tilde{m}-j}})_2 = 2^{\tilde{m}}b_{\tilde{m}-1-j}+(\overline{b_{\tilde{m}-2-j}\ldots b_0 b_{\tilde{m}-1} b_{\tilde{m}-j}0})_2\\
&\equiv \beta_{j+1}\pmod{2^{\tilde{m}}-1}\Rightarrow\ \beta_j\equiv 2^j\beta\pmod{2^{\tilde{m}}-1}\Rightarrow\ \alpha^{\beta_j}=\alpha^{2^j\beta}
\end{aligned} \qquad (۴-۲۵)$$

که نشان می‌دهد $\{\alpha^{\beta_j}\}_j$ مجموعه مزدوج‌های α^β است. برای این که نشان دهیم $\alpha^\beta\in\mathcal{H}_{\tilde{m}}^{(l)}$، باید ثابت کنیم که تمام مزدوج‌های آن متعلق به $\mathcal{G}_{\tilde{m}}^{(l)}$ هستند و یا به طور معادل، باید نشان دهیم $0\le\beta_j\le 2^{\tilde{m}-1}+2^l-1$. واضح است که $\beta_j>0$؛ برای اثبات طرف دیگر نامساوی فوق، دو حالت در نظر می‌گیریم:

۱. با ارزش‌ترین بیت β_j صفر است؛

$$b_{\tilde{m}-1-j}=0\Rightarrow\beta_j<2^{\tilde{m}-1}<2^{\tilde{m}-1}+2^l-1 \qquad (۴-۲۶)$$

۲. با ارزش‌ترین بیت β_j یک است؛ با توجه به خاصیت قرارگیری ۱ها در دنباله، $\tilde{m}-l-۱$ بیت باارزش پس از باارزش‌ترین بیت، همگی صفر هستند:

$$b_{\tilde{m}-۱-j}=۱ \quad \Rightarrow \quad b_{\tilde{m}-۲-j}=\cdots=b_{l-j}=۰ \Rightarrow \quad \beta_j \le ۲^{\tilde{m}-۱}+\sum_{j=۰}^{l-۱} ۲^j$$
$$\Rightarrow \quad \beta_j \le ۲^{\tilde{m}-۱}+۲^l-۱ \qquad (۲۷\text{–}۴)$$

تاکنون نشان دادیم که دنباله‌های دودویی با نحوه‌ی مناسب قرارگیری ۱ها، متناظر با ریشه‌هایی مجزا در $h(x)$ هستند. برای تکمیل اثبات باید نشان دهیم تمام ریشه‌های $h(x)$ به این صورت پوشانده شده‌اند. دقت کنید که اگر بسط $\tilde{m}$رقمی مبنای ۲ عدد اعشاری β خاصیت ذکر شده را در مورد نحوه قرارگیری ۰ و ۱ها در کنار هم نداشته باشد، حداقل یکی از β_jها از $۲^{\tilde{m}-۱}+۲^l-۱$ بزرگتر خواهد بود . در نتیجه تمام مزدوج‌های α^β در $\mathcal{G}_{\tilde{m}}^{(l)}$ قرار نمی‌گیرند. ■

قضیه ۴–۲ پارامتر $\tilde{k}$ ایجاد شده در این نحوه طراحی کد را به یک مساله ترکیباتی تبدیل می‌کند. در پیوست الف نشان می‌دهیم که $|\mathcal{H}_{\tilde{m}}^{(l)}| \gtrapprox \mathcal{O}\left(۲^{(l+۱)\frac{\ln \tilde{m}-l-۱}{\tilde{m}-l-۱}}\right)$.

۴–۳–۲ الگوریتم تولید ماتریس

با توجه به قضیه ۴–۱ به هر نحوی که از هریک از زوج‌های مکمل، یک عضو را انتخاب کنیم می‌توان به ماتریسی با ضریب همدوسی مورد نظر دست یافت. از آن‌جا که $۲^{\tilde{k}-۱}$ زوج مکمل وجود دارد، جدا از نحوه چینش بردارهای کد به عنوان ستون‌های ماتریس، $۲^{۲^{\tilde{k}-۱}}$ حالت مختلف برای ساختن ماتریس حسگر وجود دارد. از بین این تعداد حالت بسیار زیاد، ویژگی برخی از آن‌ها متمایز کننده است. چنانچه بتوانیم انتخاب عضو از زوج‌های مکمل را به نحوی انجام دهیم که بردارهای انتخاب شده نسبت به چرخش دوری بسته باشند، می‌توان پیچیدگی محاسباتی در الگوریتم بازسازی بردار تنک از روی نمونه‌ها را کاهش داد (این مطلب در بخش ۴–۶ مفصلاً شرح داده خواهد شد). یادآوری می‌شود که کدهای BCH، خود زیرمجموعه‌ای از کدهای گردشی هستند. بنابراین اگر در انتخاب عضو از زوج‌های مکمل با دقت عمل کنیم، می‌توانیم خاصیت گردشی کد را کماکان حفظ کنیم. از آن‌جا که طول کد ($\tilde{n}=۲^{\tilde{m}}-۱$) عددی فرد است، در هر زوج مکمل دقیقاً یکی از بردارها تعداد زوجی المان ۱ دارد. اکنون اگر تمام بردارهای کد با وزن فرد (و یا زوج) را دور بریزیم، از هر زوج مکمل دقیقاً یکی را انتخاب کرده‌ایم و همچنین گردش‌های دوری بردارهای کد باقیمانده نیز حفظ شده‌اند. مراحل ساخت

ماتریس حسگر به طور خلاصه به شرح زیر است:

۱. برای یک مقدار k (مرتبه RIP) داده شده، قرار دهید $i = \lceil \log_۲(k) \rceil$ و انتخاب کنید $\tilde{m} \geq i$. ماتریس نهایی $m = ۲^{\tilde{m}} - ۱$ سطر خواهد داشت؛

۲. فرض کنید $\mathcal{H}_{seq}$ مجموعه تمام دنباله‌های دودویی به طول $\tilde{m}$ باشد که بین هر دو ۱ حداقل i صفر به صورت دوری قرار گرفته باشد. همچنین فرض کنید $\mathcal{H}_{dec}$ نمایش اعشاری این دنباله‌ها باشد؛

۳. یک ریشه اولیه دلخواه از میدان $GF(۲^{\tilde{m}})$ مثل α انتخاب کنید و قرار دهید:

$$\mathcal{H} = \{\alpha^r \mid r \in \mathcal{H}_{dec}\} \tag{۲۸-۴}$$

۴. چند جمله ای‌های مولد کد و آزمون توازن را به صورت زیر تعریف کنید:

$$\begin{aligned} h(x) &= \prod_{r \in \mathcal{H}} (x - r) \\ g(x) &= \frac{x^{۲^{\tilde{m}}-۱} + ۱}{h(x)} \end{aligned} \tag{۲۹-۴}$$

۵. ماتریس $\tilde{\mathbf{A}}_{(۲^{\tilde{m}}-۱)\times(۲^{\deg(h)-۱})}$ را با کنار هم قرار دادن بردارهای کد با توازن زوج تشکیل دهید (ترتیب قرارگرفتن بردارها دلخواه است)؛

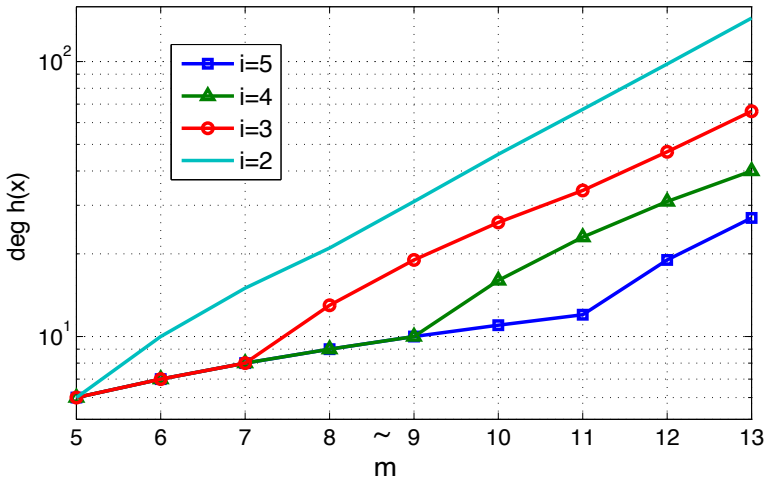
۶. صفرهای ماتریس $\tilde{\mathbf{A}}$ را با ۱- جایگذاری و ستون‌ها را با ضرب کردن در $\frac{۱}{\sqrt{m}}$ یکه کنید تا ماتریس $\mathbf{A}_{(۲^{\tilde{m}}-۱)\times ۲^{\deg h-۱}}$ بدست آید؛

. به عنوان یک مثال ساده، حالت $\tilde{m} = i$ را بررسی می‌کنیم؛ به راحتی می‌توان نشان داد که تعداد ۱ها در دنباله‌های دودویی حداکثر برابر با ۱ است، در نتیجه $\mathcal{H}_{dec} = \{۰, ۲^۰, ۲^۱, ۲^۲, \ldots, ۲^{۲^{i-۱}}\}$. این به آن معنی است که $h(x)$ از ضرب $x + ۱$ در چندجمله‌ای مینیمال α حاصل می‌شود. از آن‌جا که برای تولید کد از چندجمله‌ای $g(x)$ به جای $(x + ۱)g(x)$ استفاده می‌کنیم (دور ریختن توازن‌های فرد)، $h(x)$ موثر همان چندجمله‌ای مینیمال α خواهد بود که یک چندجمله‌ای اولیه است. در این حالت ماتریس $\tilde{\mathbf{A}}$ یک ماتریس مربعی $(۲^i - ۱) \times (۲^i - ۱)$ است که ستون‌های آن از گردش‌های دوری یک دنباله شبه تصادفی[۱۲] حاصل شده‌اند و ضریب همدوسی آن برابر با $\frac{۱}{۲^i - ۱}$ است.

[۱۲] Pseudo Noise Sequence (PNS)

$\tilde{m}$	$h(x)$
۴	$x^{۵}+x^{۴}+x^{۲}+۱$
۶	$x^{۷}+x^{۶}+x^{۲}+۱$
۸	$x^{۱۳}+x^{۱۲}+x^{۱۰}+x^{۹}+x^{۸}+x^{۴}+x^{۳}+۱$
۱۰	$x^{۲۶}+x^{۲۵}+x^{۲۴}+x^{۲۰}+x^{۱۶}+x^{۱۴}+x^{۱۳}+x^{۱۲}$ $+x^{۱۰}+x^{۹}+x^{۷}+x^{۵}+x^{۴}+x^{۳}+x+۱$

جدول ۴-۱: چندجمله‌ای آزمون توازن برای مقادیر متفاوت $\tilde{m}$ و $i=۳$.



شکل ۴-۱: درجه چندجمله‌ای $h(x)$ برای مقادیر متفاوت $\tilde{m}$ و i.

جدول ۴-۱ چند جمله ای‌های آزمون توازن برای حالت $i=۳$ (و نتیجتاً $k<۸$) را نمایش می‌دهد. همچنین شکل ۴-۱ درجه $h(x)$ را برای چند انتخاب مختلف $\tilde{m}$ و i نشان می‌دهد؛ نرخ رشد این درجه بر حسب $\tilde{m}$ در مقادیر کوچک خطی و از جایی به بعد، نمایی است.

۴-۴ ماتریس‌های مختلط به کمک کدهای غیردودویی

در بخش قبل با استفاده از کدهای دودویی، ماتریس‌های حسگر دو قطبی ساختیم. در این بخش به دنبال تعمیم روش قبلی به کدهای p-سمبلی[۱۳] هستیم. برای این منظور دو مانع اصلی در سر راه است: ۱) $\tilde{d}_{min}$ در کدهای p-سمبلی، تنها تعداد مکان‌های نابرابر در دو کد را نمایش می‌دهد و اطلاعات بیشتری در مورد ارتباط مقادیر نابرابر در اختیار قرار نمی‌دهد (در حالت دودویی، نابرابربودن دو مقدار به منزله ۰ و ۱ بودن آن‌هاست)؛ ۲) در

[۱۳] p-ary Codes

کدهای دودویی با تبدیل ۰ به ۱−، قدر مطلق درایه‌ها را برابر کردیم اما در کدهای p–سمبلی، p مقدار با اندازه‌های برابر مورد نیاز است. برای حل مشکل دوم، از به توان رساندن[۱۴] کد بهره می‌جوییم که منجر به تولید ماتریس مختلط می‌شود.

فرض کنید $\mathcal{C}(\tilde{n},\tilde{k};p)$ یک کد خطی p–سمبلی روی میدان $GF(p)$ (p توانی از یک عدد اول است) با حداقل فاصله $\tilde{d}_{min}$ باشد به نحوی که بردار تمام یک $\mathbf{1}_{\tilde{n}\times 1}$ متعلق به فضای کد باشد؛ به دلیل خطی بودن کد، تمامی بردارهای $\mathbf{0}_{\tilde{n}\times 1}, \mathbf{1}_{\tilde{n}\times 1}, \ldots, (p-1)_{\tilde{n}\times 1}$ نیز در فضای کد قرار دارند. مشابه حالت دودویی، برای هر دو بردار کد $\mathbf{a}_{\tilde{n}\times 1}$ و $\mathbf{b}_{\tilde{n}\times 1}$ اگر تعریف کنیم $\mathbf{c}_{\tilde{n}\times 1} \triangleq \mathbf{a} \oplus -\mathbf{b}$ که $\oplus$ جمع المان به المان به پیمانه p است، یکی از دو حالت زیر رخ می‌دهد:

۱. $\mathbf{c} = \mathbf{0}_{\tilde{n}\times 1}$ یا $\mathbf{c} = \mathbf{1}_{\tilde{n}\times 1}$ یا ... یا $\mathbf{c} = (p-1)_{\tilde{n}\times 1}$؛

۲. $\mathbf{c} \notin \{\mathbf{0}_{\tilde{n}\times 1}, \mathbf{1}_{\tilde{n}\times 1}, \ldots, (p-1)_{\tilde{n}\times 1}\}$؛ بنابراین

$$\begin{cases} d(\mathbf{c}_{\tilde{n}\times 1}, \mathbf{0}_{\tilde{n}\times 1}) \geq \tilde{d}_{min} \\ d(\mathbf{c}_{\tilde{n}\times 1}, \mathbf{1}_{\tilde{n}\times 1}) \geq \tilde{d}_{min} \\ \vdots \\ d(\mathbf{c}_{\tilde{n}\times 1}, (p-1)_{\tilde{n}\times 1}) \geq \tilde{d}_{min} \end{cases} \qquad (۳۰-۴)$$

که نشان می‌دهد $\mathbf{c}_{\tilde{n}\times 1}$ حداکثر $\tilde{n} - \tilde{d}_{min}$ عضو از هریک از اعضای مجموعه $\{0, 1, \ldots, p-1\}$ را در بر می‌گیرد. برای هر $0 \leq i \leq p-1$، N_i را برابر تعداد تکرارهای عدد i در بردار $\mathbf{c}_{\tilde{n}\times 1}$ تعریف می‌کنیم. در نتیجه داریم $N_i \leq \tilde{n} - \tilde{d}_{min}$ و نیز $\sum_{i=0}^{p-1} N_i = \tilde{N}$ که نتیجه می‌دهد:

$$N_i = \tilde{n} - \sum_{j\neq i} N_j \geq \tilde{n} - (p-1)(\tilde{n} - \tilde{d}_{min}) \qquad (۳۱-۴)$$

بنابراین

$$\underbrace{\tilde{n} - (p-1)(\tilde{n} - \tilde{d}_{min})}_{N_{min}} \leq N_i \leq \underbrace{\tilde{n} - \tilde{d}_{min}}_{N_{max}}, \qquad (۳۲-۴)$$

که معادل است با :

$$\left|N_i - \frac{N_{min} + N_{max}}{2}\right| \leq \frac{N_{max} - N_{min}}{2} \qquad (۳۳-۴)$$

[۱۴] Exponentiation

به جای جفت‌کردن بردارهای کد در حالت دودویی (جفت‌های مکمل) در این‌جا می‌توان بردارهای کد را به مجموعه‌هایی به شکل $\{\mathbf{a}, \mathbf{a} \oplus \mathbf{1}_{\tilde{n}\times 1}, \ldots, \mathbf{a} \oplus (p-1)_{\tilde{n}\times 1}\}$ افراز کرد. در حقیقت، این افراز، تقسیم گروه کل بردارهای کد به زیر گروه $\{\mathbf{0}_{\tilde{n}\times 1}, \ldots, (p-1)_{\tilde{n}\times 1}\}$ را نشان می‌دهد (گروه نسبت به عمل $\oplus$).

قضیه ۴-۳ فرض کنید $\mathcal{C}(\tilde{n}, \tilde{k}; p)$ یک کد خطی p-سمبلی بر روی میدان $GF(p)$ با $\tilde{d}_{min}$ باشد به طوری که بردار تمام ۱ متعلق به فضای کد باشد؛ همچنین فرض کنید $\tilde{\mathbf{A}}_{\tilde{n}\times p^{\tilde{k}-1}}$ ماتریس حاصل از کنار هم قراردادن دسته‌ای از بردارهای کد باشد به نحوی که از هر مجموعه به شکل $\{\mathbf{a}, \mathbf{a} \oplus \mathbf{1}_{\tilde{n}\times 1}, \ldots, \mathbf{a} \oplus (p-1)_{\tilde{n}\times 1}\}$ دقیقاً یک عضو انتخاب شده باشد. اکنون اگر ماتریس $\mathbf{A}_{\tilde{n}\times p^{\tilde{k}-1}}$ را از به توان رساندن و سپس یکه کردن ماتریس $\tilde{\mathbf{A}}$ به صورت زیر بسازیم:

$$\tilde{\mathbf{A}} = [\tilde{a}_{\alpha\beta}]_{\alpha,\beta} \Rightarrow \mathbf{A} = \frac{1}{\sqrt{\tilde{n}}}\left[e^{j\frac{2\pi}{p}\tilde{a}_{\alpha\beta}}\right]_{\alpha,\beta} \qquad (۳۴-۴)$$

ضریب همدوسی ماتریس $\mathbf{A}$ حداکثر برابر با $\frac{p(p-1)\tilde{n}-p^2\tilde{d}_{min}}{2\tilde{n}}$ است.

اثبات: ابتدا دقت کنید که ستون‌های $\mathbf{A}$ یکه هستند:

$$\|\mathbf{a}_\beta\| = \left\| \frac{1}{\sqrt{\tilde{n}}}[e^{j\frac{2\pi}{p}\tilde{a}_{1,\beta}} \ \ldots \ e^{j\frac{2\pi}{p}\tilde{a}_{\tilde{n},\beta}}]^T \right\| = 1 \qquad (۳۵-۴)$$

فرض کنید $\mathbf{a}_\alpha, \mathbf{a}_\beta$ ستون‌های متمایزی از ماتریس $\mathbf{A}$ و $\tilde{\mathbf{a}}_\alpha, \tilde{\mathbf{a}}_\beta$ ستون‌های متناظر در ماتریس $\tilde{\mathbf{A}}$ باشند و تعریف کنید $\mathbf{c} = \tilde{\mathbf{a}}_\alpha \oplus -\tilde{\mathbf{a}}_\beta$. همچنین مشابه قبل، فرض کنید N_i بیانگر تعداد تکرارهای عدد i در بردار $\mathbf{c}$ باشد $(0 \le i \le p-1)$. داریم:

$$|\langle \mathbf{a}_\alpha, \mathbf{a}_\beta \rangle| = |\tilde{\mathbf{a}}_\beta^H . \tilde{\mathbf{a}}_\alpha| = \frac{1}{\tilde{n}}\left|\sum_{i=1}^{\tilde{n}} e^{j\frac{2\pi}{p}(\tilde{a}_{i,\alpha}-\tilde{a}_{i,\beta})}\right| = \frac{\left|\sum_{i=1}^{\tilde{n}} e^{j\frac{2\pi}{p}c_i}\right|}{\tilde{n}} = \frac{\left|\sum_{i=0}^{p-1} N_i e^{j\frac{2\pi}{p}i}\right|}{\tilde{n}} \qquad (۳۶-۴)$$

از آن‌جا که $e^{j\frac{2\pi}{p}}$ یک ریشه چندجمله‌ای $1 + x + \cdots + x^{p-1}$ است، برای تمام مقادیر γ داریم:

$$\left|\sum_{i=0}^{p-1} N_i e^{j\frac{2\pi}{p}i}\right| = \left|\sum_{i=0}^{p-1} (N_i - \gamma) e^{j\frac{2\pi}{p}i}\right| \le \sum_{i=0}^{p-1} |N_i - \gamma| \qquad (۳۷-۴)$$

با استفاده از نامساوی‌های (۴-۳۲) و (۴-۳۳) و قرار دادن $\gamma = \frac{N_{min}+N_{max}}{2}$ خواهیم داشت:

$$\left|\sum_{i=0}^{p-1} N_i e^{j\frac{2\pi}{p}i}\right| \le p\ \frac{N_{max} - N_{min}}{2} = \frac{p(p-1)\tilde{n} - p^2\tilde{d}_{min}}{2} \qquad (۳۸-۴)$$

که کران بالای مورد نظر برای ضریب همدوسی ماتریس $\mathbf{A}$ را نتیجه می‌دهد:

$$|\langle \mathbf{a}_\alpha, \mathbf{a}_\beta \rangle| \leq \frac{p(p-1)\tilde{n} - p^2 \tilde{d}_{min}}{2\tilde{n}} \qquad (۳۹-۴)$$

■

نکته ۱ بهترین انتخاب γ در (۴-۳۷) که کوچکترین کران بالا را ارائه می‌دهد، میانه N_iهاست و نه لزوماً مقدار به کار رفته در (۴-۳۸). اما از آن‌جا که رابطه مشخصی برای میانه این اعداد وجود ندارد، از مقدار یاد شده برای γ استفاده می‌شود.

نکته ۲ برای برقراری RIP مرتبه k، با استفاده از نامساوی یاد شده بر روی ضریب همدوسی در قضیه ۴-۳باید داشته باشیم:

$$\frac{\tilde{d}_{min}}{\tilde{n}} > \frac{p-1}{p} - \frac{2}{p^2(k-1)} \geq \frac{p-1}{p}(1 - \frac{8}{kp^2}) \qquad (۴۰-۴)$$

در نتیجه $\tilde{d}_{min}$ باید بسیار نزدیک به $\frac{p-1}{p}\tilde{n}$ باشد؛ به عبارت بهتر، برای مقادیر بزرگ p، $\tilde{d}_{min}$ باید تقریباً برابر با $\tilde{n}$ باشد. در ادامه، مشابه حالت دودویی، وجود چنین کدهایی را به کمک ساختار BCH اثبات می‌کنیم.

۴-۴-۱ کدهای BCH p-سمبلی با $\tilde{d}_{min}$ بزرگ

کدهای BCH p-سمبلی بر روی میدان $GF(p)$ تعریف می‌شوند و طول بردارهای کد به صورت $\tilde{n} = p^{\tilde{m}} - 1$ انتخاب می‌شود و بردارهای کد توسط چندجمله‌ای مولد کد $g(x)$ متعلق به $GF(p)[x]$ تولید می‌شوند. مشابه حالت دودویی، تمام ریشه‌های $g(x)$ و چندجمله‌ای آزمون توازن $h(x)$ در $GF(p^{\tilde{m}})$ قرار دارند. از آن‌جا که:

$$\prod_{\substack{r \in GF(p^{\tilde{m}}) \\ r \neq 0}} (x - r) = x^{p^{\tilde{m}}-1} - 1 \qquad (۴۱-۴)$$

در حالت p-سمبلی داریم:

$$h(x) = \frac{x^{p^{\tilde{m}}-1} - 1}{g(x)} \qquad (۴۲-۴)$$

در این حالت نیز می‌توان مشابه حالت دودویی، قضیه‌ای بر روی حداقل فاصله کد به صورت زیر بیان کرد: اگر α یک ریشه اولیه میدان $GF(p^{\tilde{m}})$ و $\{\alpha^{i_1}, \ldots, \alpha^{i_d}\}$ زیرمجموعه‌ای از ریشه‌های متمایز $g(x)$ باشد به طوری که

$i_1,\ldots,i_d$ یک تصاعد حسابی تشکیل دهند، داریم $\tilde{d}_{min} \geq d+1$ [۶۶]. برای اثبات، فرض کنید که $[c_1,\ldots,c_{\tilde{n}}]^T$ یک بردار کد ناصفر باشد؛ در نتیجه $g(x)|\sum_{j=1}^{\tilde{n}} c_j x^{j-1}$[۱۵] ولذا:

$$\underbrace{\begin{bmatrix} \alpha^{0\times i_1} & \alpha^{1\times i_1} & \ldots & \alpha^{(\tilde{n}-1)\times i_1} \\ \alpha^{0\times i_2} & \alpha^{1\times i_2} & \ldots & \alpha^{(\tilde{n}-1)\times i_2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{0\times i_d} & \alpha^{1\times i_d} & \ldots & \alpha^{(\tilde{n}-1)\times i_d} \end{bmatrix}}_{\mathbf{H}_{d\times\tilde{n}}} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{\tilde{n}} \end{bmatrix} = 0_{d\times 1} \qquad (۴۳-۴)$$

از آن‌جا که $\{i_1,\ldots,i_d\}$ یک تصاعد حسابی تشکیل می‌دهند، هر زیر ماتریس $d\times d$ از ماتریس $\mathbf{H}$ یک ماتریس واندرموند[۱۶] است. از این رو، هر d ستون از ماتریس $\mathbf{H}$ مستقل خطی هستند که وجود حداقل $d+1$ عضو ناصفر در بین $[c_1,\ldots,c_{\tilde{n}}]^T$ را ایجاب می‌کند (کران مورد نظر بر روی حداقل فاصله).

فرض کنید چندجمله‌ای مولد کد $g(x)$ طوری انتخاب شده باشد که مجموعه $\{\alpha^{p^{\tilde{m}-1}+\frac{p^l-1}{p-1}+1}, \alpha^{p^{\tilde{m}-1}+\frac{p^l-1}{p-1}+2},\ldots,\alpha^{p^{\tilde{m}}-2}\}$ برای $l<\tilde{m}$ زیرمجموعه‌ای از ریشه‌های $g(x)$ باشد. با توجه به نکته بیان شده در مورد ارتباط بین حداقل فاصله کد و طول تصاعد حسابی موجود در بین توان‌های α میان ریشه‌های $g(x)$، داریم:

$$\begin{aligned} \tilde{d}_{min} &\geq p^{\tilde{m}} - p^{\tilde{m}-1} - \frac{p^l-1}{p-1} - 1 = (p^{\tilde{m}}-1)\left(1 - \frac{p^{\tilde{m}-1}}{p^{\tilde{m}}-1} - \frac{p^l-1}{(p^{\tilde{m}}-1)(p-1)}\right) \\ &= \tilde{n}\left(\frac{p-1}{p} - \frac{p^{l+1}-1}{p(p-1)(p^{\tilde{m}}-1)}\right) \\ &\Rightarrow \frac{\tilde{d}_{min}}{\tilde{n}} \geq \frac{p-1}{p}\left(1 - \frac{p^{l+1}-1}{(p-1)^2(p^{\tilde{m}}-1)}\right) \end{aligned} \qquad (۴۴-۴)$$

که نشان میدهد چنین کدی شرط مطلوب برای حداقل فاصله را دارد. برای بدست آوردن $g(x)$، مشابه حالت دودویی تعریف کنید:

$$\begin{aligned} \mathcal{G}_{\tilde{m}}^{(l)} &= \{\alpha^0, \alpha^1, \ldots, \alpha^{p^{\tilde{m}-1}+\frac{p^l-1}{p-1}}\} \\ \mathcal{H}_{\tilde{m}}^{(l)} &= \{r \in \mathcal{G}_{\tilde{m}}^{(l)} \mid \forall\, j\in\mathbb{N}:\ r^{p^j} \in \mathcal{G}_{\tilde{m}}^{(l)}\} \end{aligned} \qquad (۴۵-۴)$$

به بیان ساده‌تر، $\mathcal{H}_{\tilde{m}}^{(l)}$ زیرمجموعه‌ای از $\mathcal{G}_{\tilde{m}}^{(l)}$ است که تمامی مزدوج‌های هر عضو آن نسبت به $GF(p)$ را نیز

[۱۵] $x|y$ یعنی y بر x بخش پذیر است.

[۱۶] Vandermonde

شامل می‌شود. در نتیجه:

$$h(x) \triangleq \prod_{h \in \mathcal{H}_{\tilde{m}}^{(l)}} (x - \alpha^h) \in GF(p)[x] \tag{۴۶-۴}$$

چندجمله‌ای $g(x)$ را نیز می‌توان از روی $h(x)$ تعریف شده محاسبه کرد. نحوه طراحی چند جمله ای‌های $g(x), h(x)$ نشان می‌دهد که کد حاصل، حداقل فاصله مطلوب را دارد. اما این نکته که بردار تمام ۱ متعلق به فضای کد است، هنوز چندان روشن نیست. دقت کنید که $\alpha^{\circ} = ۱ \in \mathcal{H}_{\tilde{m}}^{(l)}$ ریشه $h(x)$ است و در نتیجه نسبت به $g(x)$ اول است. پس:

$$\begin{cases} g(x) \big| x^{\tilde{n}} - ۱ = (x-۱)(۱ + x + \cdots + x^{\tilde{n}-۱}) \\ \\ gcd\big(g(x)\ ,\ x-۱\big) = ۱ \end{cases} \Rightarrow g(x) \big| ۱ + x + \cdots + x^{\tilde{n}-۱} \tag{۴۷-۴}$$

که تعلق $۱_{\tilde{n}\times ۱}$ به فضای کد را نشان می‌دهد. مجدداً در این ساختار $\tilde{n}$ و $\tilde{d}_{min}$ (کران پایین) معلوم هستند در حالی که $\tilde{k}$ هنوز مشخص نشده است. قضیه زیر تعداد اعضای $\mathcal{H}_{\tilde{m}}^{(l)}$ و در نتیجه درجه $h(x)$ را به صورت یک مسئله ترکیباتی بیان می‌کند:

قضیه ۴-۴ $|\mathcal{H}_{\tilde{m}}^{(l)}|$ برابر است با تعداد دنباله‌های دودویی به طول $\tilde{m}$ به طوری که بین هر دو ۱ متوالی دست کم $\tilde{m} - l - ۱$ صفر به صورت گردشی وجود داشته باشد.

اثبات: روند کاملاً مشابه اثبات قضیه ۴-۲ است با این تفاوت که در این حالت به جای بسط مبنای ۲، بسط مبنای p استفاده می‌شود. ■

در پیوست الف نشان می‌دهیم $|\mathcal{H}_{\tilde{m}}^{(l)}| = \mathcal{O}(\gamma^{l+۱})$ که γ بزرگترین ریشه $x^{\tilde{m}-l-۱} - x - ۱$ است. در این صورت، ماتریس حسگر $m \times n$ ساخته شده، شرط RIP از مرتبه k را ارضا خواهد کرد که:

$$\begin{cases} m & = & p^{\tilde{m}} - ۱ \\ \log_p n & = & |\mathcal{H}_{\tilde{m}}^{(l)}| = \mathcal{O}\big(\gamma^{l+۱}\big) \\ k_{max} & \geq & ۲\frac{p-۱}{p} p^{\tilde{m}-l-۱} \geq p^{\tilde{m}-l-۱} \end{cases} \tag{۴۸-۴}$$

که منظور از k_{max} حداکثر مرتبه RIP است که با استفاده از استدلالات مبتنی بر ضریب همدوسی، قابل تضمین است. با استفاده از نامساوی $\ln \gamma \geq \frac{\ln(\tilde{m}-l-۱)}{\tilde{m}-l-۱}$ که در پیوست الف نشان داده شده است، می‌توان ثابت کرد $\gamma^{\frac{\log_p k_{max}}{\log_p \log_p k_{max}}} \geq p$. در نتیجه برای ماتریس‌های حسگر ساخته شده داریم:

$$m \leq \mathcal{O}\left(k_{max} (\log_p n)^{\frac{\log_p k_{max}}{\log_p \log_p k_{max}}} \right) \tag{۴۹-۴}$$

جدول ۴-۲: $\frac{\mu_{BCH}}{\mu_{WB}}$ برای چند ماتریس p–سمبلی $(p^{۲l}-۱)\times p^{۳l}$ با pهای مختلف.

	$p=۲$	$p=۳$	$p=۵$	$p=۷$
$l=۱$	—	۱٫۱۸۶۳	۱٫۱۰۰۹	۱٫۰۷۰۹
$l=۲$	۱٫۱۲۹۶	۱٫۰۵۴۹	۱٫۰۱۹۸	۱٫۰۱۰۲
$l=۳$	۱٫۰۶۱۸	۱٫۰۱۸۴	۱٫۰۰۴۰	۱٫۰۰۱۵

به وضوح، کران بالا بر روی m در ماتریس‌های حسگر تصادفی، $m \leq \mathcal{O}(k_{max}\log_p n)$، به مراتب کوچکتر از کران بالای بدست آمده برای این ماتریس‌ها در رابطه (۴-۴۹) است. این ضعف ناشی از استفاده از استدلالات مبتنی بر ضریب همدوسی است که شرطی قوی‌تر از RIP است. همان طور که قبلاً اشاره شد، نامساوی ولش[۱۷] کران پایینی برای ضریب همدوسی یک ماتریس نسبت به ابعاد آن ارائه می‌دهد. از آن‌جا که برقراری شرط RIP در ماتریس‌های معرفی شده توسط ضریب همدوسی صورت گرفته است، به جای مقایسه ابعاد این ماتریس‌ها با ماتریس‌های تصادفی که شرط RIP را با مرتبه یکسانی ارضا می‌کنند، مطلوب‌تر آن است که ضریب همدوسی بدست آمده را با کران ولش مقایسه کنیم. برای این منظور در جدول ۴-۲ نسبت ضریب همدوسی‌های ماتریس‌های بدست آمده به ضریب همدوسی پیش بینی شده توسط نامساوی ولش ($\frac{\mu_{BCH}}{\mu_{WB}}$) برای چند حالت مختلف محاسبه شده است. برای این مقایسه، از حالت خاص $\tilde{m}=۲l$ استفاده کرده‌ایم که منجر به تولید ماتریس‌های $(p^{۲l}-۱)\times p^{۳l}$ می‌شود. به راحتی می‌توان نشان داد که برای هر p، $\mathcal{H}_{۲l}^{(l)}=\{۰\}\cup\{p^i\}_{i=۰}^{۲l-۱}\cup\{p^{l+i}+p^i\}_{i=۰}^{l-۱}$. نتایج جدول ۴-۲ نشان می‌دهند که با افزایش p، ضریب همدوسی ماتریس‌های بدست آمده به کران ولش نزدیک‌تر می‌شوند.

۴-۴-۲ الگوریتم تولید ماتریس

در قضیه ۴-۳ به این نکته اشاره شد که به منظور دست‌یابی به یک ماتریس حسگر مطلوب باید از هر زیرمجموعه به شکل $\{\mathbf{a}, \mathbf{a}\oplus ۱_{\tilde{n}\times ۱}, \ldots, \mathbf{a}\oplus(p-۱)_{\tilde{n}\times ۱}\}$ در فضای کد، دقیقاً یک بردار کد انتخاب شود. در حالت دودویی ($p=۲$) دیدیم که این انتخاب را می‌توان به نحوی انجام داد که بردارهای انتخاب شده کماکان خاصیت چرخش دوری خود را حفظ کنند (مثلاً با انتخاب تمام بردارهای کد با وزن زوج). مشابه این عمل را نیز می‌توان در حالت p–سمبلی به کار برد. از ان جا که طول بردار کد ($\tilde{n}$) نسبت به اندازه میدان (p) اول است، $x=۱$ ریشه دقیقاً یکی از چند جمله ای‌های متناظر با بردارهای $\{\mathbf{a}, \mathbf{a}\oplus ۱_{\tilde{n}\times ۱}, \ldots, \mathbf{a}\oplus(p-۱)_{\tilde{n}\times ۱}\}$ است (تعمیم وزن زوج

[۱۷] Welch

به حالت p–سمبلی). در نتیجه اگر $x = ۱$ به ریشه‌های $g(x)$ اضافه شود، به طور خودکار از هر مجموعه به شکل یاد شده، دقیقاً یک عضو انتخاب می‌شود و کد حاصل همچنان یک کد گردشی است. از این رو، می‌توان مراحل الگوریتم تولید ماتریس را برای p داده‌شده به صورت توانی از یک عدد اول توسط گام‌های زیر بیان کرد:

۱. عدد طبیعی $\tilde{m}$ را انتخاب کنید و قرار دهید $m = p^{\tilde{m}} - ۱$؛

۲. عدد صحیح $۰ \leq l \leq \tilde{m} - ۱$ را اختیار کنید؛ ماتریس نهایی، شرط RIP با مرتبه $k < ۲\frac{p-۱}{p}\frac{p^{\tilde{m}}-۱}{p^{l+۱}-۱} + ۱$ را ارضا خواهد کرد؛

۳. مجموعه $\mathcal{H}_{seq}^{(\tilde{m},l)}$ شامل تمام دنباله‌های دودویی به طول $\tilde{m}$ را که میان هر دو یک متوالی دست کم $\tilde{m} - l - ۱$ صفر به صورت گردشی وجود دارد، تشکیل دهید. همچنین $\mathcal{H}_{\tilde{m}}^{(l)}$ شامل تمام اعداد اعشاری است که بسط مبنای p آن‌ها عضوی از مجموعه $\mathcal{H}_{seq}^{(\tilde{m},l)}$ باشد؛

۴. α را برابر یا یک ریشه اولیه میدان $GF(p^{\tilde{m}})$ قرار دهید و تعریف کنید:

$$h(x) = \prod_{r \in \mathcal{H}_{\tilde{m}}^{(l)} - \{۰\}} (x - \alpha^r)$$

همچنین قرار دهید $n = p^{|\mathcal{H}_{\tilde{m}}^{(l)}|-۱}$ ؛

۵. تمام بردارهای کد تعریف شده توسط چندجمله‌ای آزمون توازن $h(x)$ و چندجمله‌ای مولد کد $g(x) = \frac{x^{p^{\tilde{m}}-۱}-۱}{h(x)}$ را به عنوان ستون‌های ماتریس $\tilde{\mathbf{A}}_{m\times n}$ قرار دهید؛

۶. ماتریس حسگر نهایی را به صورت

$$\mathbf{A}_{m\times n} = \frac{۱}{\sqrt{m}}\left[e^{j۲\pi\frac{\tilde{a}_{i,j}}{p}}\right]$$

تعریف کنید که در آن $\tilde{a}_{i,j}$ها المان‌های ماتریس $\tilde{\mathbf{A}}$ هستند.

۴-۵ ادغام ماتریس‌ها

روش‌های طراحی ماتریس حسگر یقینی تقریباً در تمام موارد بر اساس کمینه‌سازی ضریب همدوسی ماتریس و بر پایه میدان‌های متناهی بنا شده‌اند. به همین دلیل، تعداد سطرهای این ماتریس‌ها که رابطه بسیار نزدیکی با اندازه میدان دارد، به دسته خاصی از اعداد طبیعی مانند توان‌های اعداد اول محدود می‌شود. در این بخش، به

کمک دو رویکرد متفاوت با تلفیق ماتریس‌های حسگر، ماتریس جدیدی معرفی می‌کنیم که ابعاد آن حوزه وسیع‌تری از مقادیر را می‌پذیرد.

در روش نخست، با استفاده از ماتریس‌های دودویی، با ثابت نگاه‌داشتن تعداد سطرها و ضریب همدوسی، تعداد ستون‌ها را افزایش می‌دهیم حال آن که در روش دوم، با تلفیق دو ماتریس حسگر دلخواه، با ثابت نگاه‌داشتن ضریب همدوسی و یا مرتبه RIP، ماتریس جدیدی با تعداد سطر و ستون بیشتر طراحی می‌کنیم.

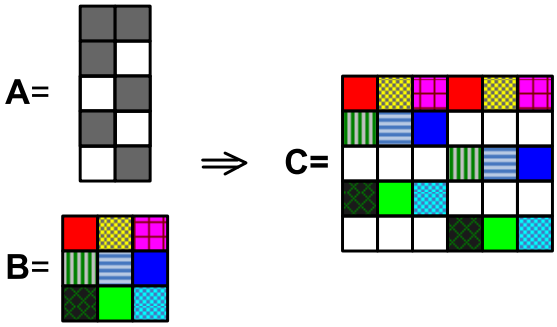
۴-۵-۱ ادغام با ماتریس‌های دودویی

طراحی ماتریس‌های دودویی با ضریب همدوسی کوچک به دلیل نامنفی بودن جملات در ضرب داخلی ستون‌ها، دشوارتر از طراحی در حالتی است که محدود به ماتریس‌های دودویی نیستیم. خوشبختانه، حداقل دو طرح دودویی موجود است: ۱) طرح Devore [۳۸] که در آن ماتریس $p^۲$ سطر دارد و هر ستون وزن p دارد (که p توانی از یک عدد اول است)؛ ۲) ماتریس‌های ساخته شده براساس کدهای متعامد نوری که در بخش ۴-۲ به آن اشاره شد. لم زیر نشان می‌دهد که چگونه می‌توان یک ماتریس دودویی و یک ماتریس غیردودویی را با یکدیگر ادغام کرد.

لم ۴-۲ فرض کنید $\mathbf{A}$ یک ماتریس دودویی با ضریب همدوسی $\mu_{\mathbf{A}} \leq \frac{۱}{k-۱}$ باشد که وزن هر ستون آن w_m است؛ همچنین فرض کنید $\mathbf{B}_{w_m \times n_۲}$ ماتریسی با درایه‌های هم‌اندازه، ستون‌های یکه و ضریب همدوسی $\mu_{\mathbf{B}} \leq \frac{۱}{k-۱}$ باشد. در این صورت، می‌توان ماتریس $\mathbf{C}_{m \times (n-۱n_۲)}$ با ستون‌های یکه و ضریب همدوسی $\mu_{\mathbf{C}} \leq \frac{۱}{k-۱}$ را با ادغام ماتریس‌های $\mathbf{A}$ و $\mathbf{B}$ ساخت.

اثبات: برای ساختن ستون lام از ماتریس $\mathbf{C}$، ابتدا $l-۱$ را به صورت $\alpha n_۲ + \beta$ که $\alpha \in \{۰, ۱, \ldots, n_۱ - ۱\}$ و $\beta \in \{۰, ۱, \ldots, n_۲ - ۱\}$ می‌نویسیم (α و β در حقیقت خارج‌قسمت و باقیمانده تقسیم $l-۱$ بر $n_۲$ هستند). فرض کنید $i_۱, \ldots, i_{w_m}$ اندیس مکان‌های ناصفر در ستون $\alpha + ۱$ ماتریس $\mathbf{A}$ باشند. حال ستون lام ماتریس را چنین تعریف می‌کنیم :

$$\begin{cases} c_{i_۱,l} &= b_{۱,\beta+۱} \\ c_{i_۲,l} &= b_{۲,\beta+۱} \\ &\vdots \\ c_{i_{w_m},l} &= b_{w_m,\beta+۱} \\ c_{s,l} &= ۰ \quad , \quad s \notin \{i_۱, \ldots, i_{w_m}\}, \end{cases} \tag{۵۰-۴}$$



شکل ۴-۲: عمل ادغام ماتریس $\mathbf{B}$ با ماتریس دودویی $\mathbf{A}$.

که $[b_{۱,\beta+۱}, \ldots, b_{w_m,\beta+۱}]^T$ ستون $\beta+۱$ام $\mathbf{B}$ است. شکل ۴-۲ به صورت نمادین این نحوه ادغام را نشان می‌دهد.

برای اثبات کران مورد نیاز بر روی ضریب همدوسی ماتریس $\mathbf{C}$، فرض کنید $\mathbf{u}_{w_m\times ۱}, \mathbf{v}_{w_m\times ۱}$ ستون‌های $l_۱$ و $l_۲$ ماتریس $\mathbf{C}$ باشند و $l_۱ - ۱ = \alpha_۱ \cdot n_۲ + \beta_۱$ و $l_۲ - ۱ = \alpha_۲ \cdot n_۲ + \beta_۲$. به راحتی می‌توان یکه بودن بردارهای $\mathbf{u}$ و $\mathbf{v}$ را بر اساس یکه بودن ستون‌های $\mathbf{B}$ نشان داد. برای بررسی ضرب داخلی میان این دو بردار، دو حالت زیر را در نظر بگیرید:

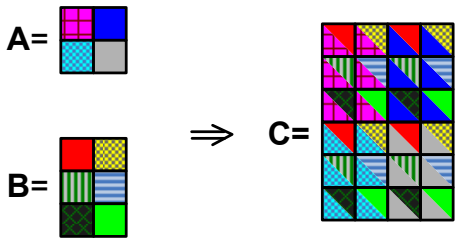
۱. هنگامی که $\alpha_۱ \neq \alpha_۲$، دو بردار $\mathbf{u}$ و $\mathbf{v}$ از بردارهای متفاوتی در ماتریس $\mathbf{A}$ حاصل شده‌اند و در نتیجه الگوهای ناصفر متفاوتی دارند. از آن‌جا که ضرب داخلی هر دو ستون ماتریس $\mathbf{A}$ کمتر از $\frac{۱}{k-۱}$ ($\frac{w_m}{k-۱}$ پیش از یکه‌کردن) است، مکان‌های ناصفر دو بردار $\mathbf{u}$ و $\mathbf{v}$ حداکثر $\frac{w_k}{k-۱}$ اشتراک دارند. به علاوه، قدرمطلق عناصر ناصفر بردارهای $\mathbf{u}$ و $\mathbf{v}$ برابر با $\frac{۱}{\sqrt{w_m}}$ است (با توجه به یکه بودن ستون‌ها و هم‌اندازه‌بودن درایه‌های ناصفر $\mathbf{B}$). در نتیجه :

$$|\langle \mathbf{u}, \mathbf{v}\rangle| = |\sum_{i=۱}^{n} u_i v_i| \leq \sum_{i=۱}^{n} |u_i v_i| = \lfloor \frac{w_m}{k-۱} \rfloor \left(\frac{۱}{\sqrt{w_m}}\right)^۲ \leq \frac{۱}{k-۱} \qquad (۵۱\text{-}۴)$$

۲. هنگامی که $\alpha_۱ = \alpha_۲$، بردارهای $\mathbf{u}$ و $\mathbf{v}$ از بردار یکسانی در ماتریس $\mathbf{A}$ ساخته شده‌اند و در نتیجه ضرب داخلی آن‌ها با ضرب داخلی بردارهای سازنده آن‌ها از $\mathbf{B}$ (ستون‌های $\beta_۱+۱$ و $\beta_۲+۱$) برابر است:

$$|\langle \mathbf{u}, \mathbf{v}\rangle| = |\langle \mathbf{b}_{\beta_۱+۱}, \mathbf{b}_{\beta_۲+۱}\rangle| \leq \frac{۱}{k-۱} \qquad (۵۲\text{-}۴)$$

بنابراین $\mathbf{C}$ ستون‌های یکه دارد و ضریب همدوسی آن کمتر از $\frac{۱}{k-۱}$ است. ■



شکل ۴-۳: ضرب کرونکر دو ماتریس ($\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$).

۴-۵-۲ ضرب کرونکر

برای ماتریس‌های دلخواه $\mathbf{A}_{m_a \times n_a}$ و $\mathbf{B}_{m_b \times n_b}$ تعریف کنید:

$$\mathbf{C}_{m_a m_b \times n_a n_b} \triangleq \mathbf{A}_{m_a \times n_a} \otimes \mathbf{B}_{m_b \times n_b} \tag{۵۳-۴}$$

که در آن $\otimes$ ضرب کرونکر[۱۸] دو ماتریس را نمایش می‌دهد، یعنی:

$$c_{\eta,\theta} = a_{\gamma,\tau} b_{\rho,\nu}, \tag{۵۴-۴}$$

که $\eta = (\gamma - ۱)m_b + \rho$ و $\theta = (\tau - ۱)n_b + \nu$ و γ, τ, ρ, ν اعداد طبیعی هستند که به ترتیب از m_a, n_a, m_b, n_b فراتر نمی‌روند. شکل ۴-۳ به صورت نمادین ضرب کرونکر دو ماتریس را نمایش می‌دهد.

لم ۴-۳ فرض کنید $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$،

(*i*) اگر $\mathbf{A}$ و $\mathbf{B}$ ستون‌های یکه داشته باشند، ستون‌های $\mathbf{C}$ نیز یکه است؛

(*ii*) $\mu_{\mathbf{C}} = \max\{\mu_{\mathbf{A}}, \mu_{\mathbf{B}}\}$؛

(*iii*) اگر $\mathbf{A}$ و $\mathbf{B}$ شرط *RIP* از مرتبه k با ثابت‌های $\delta_{k,\mathbf{A}}, \delta_{k,\mathbf{B}}$ را ارضا کنند، ماتریس $\mathbf{C}$ شرط *RIP* از مرتبه k با ثابت $\delta_{k,\mathbf{C}} \leq \delta_{k,\mathbf{A}}\delta_{k,\mathbf{B}} + \delta_{k,\mathbf{A}} + \delta_{k,\mathbf{B}}$ را ارضا می‌کند.

اثبات دو عبارت اول در [۸۰، ۶۰] آورده شده است. یک نتیجه ساده از (i) و (ii) آن است که اگر $\mathbf{A}$ و $\mathbf{B}$ ستون‌های یکه داشته باشند و ضرایب همدوسی آن‌ها به ترتیب $\mu_{\mathbf{A}} < \frac{۱}{k_{\mathbf{A}}-۱}$ و $\mu_{\mathbf{B}} < \frac{۱}{k_{\mathbf{B}}-۱}$ با $k_{\mathbf{A}}, k_{\mathbf{B}}$ صحیح باشد، ماتریس $\mathbf{C}$ شرط RIP از مرتبه $k_{\mathbf{C}} = \min\{k_{\mathbf{A}}, k_{\mathbf{B}}\}$ را ارضا می‌کند. در حقیقت عبارت (iii) که در [۴۳]

[۱۸] Kronecker

اثبات شده است، تعمیمی از این نتیجه به حالتی است که ماتریس‌های $\mathbf{A}$ و $\mathbf{B}$ شرط RIP را بدون هیچ‌گونه محدودیتی بر روی ضریب همدوسی ارضا می‌کنند.

نتیجه جالبی که می‌توان از ضرب کرونکر ماتریس‌ها بدست آورد، تولید ماتریس‌های حسگر با تعداد سطر دلخواه است. با استفاده از روش بیان شده در [۳۸]، می‌توان ماتریس‌هایی با p^k سطر تولید کرد که p یک عدد اول است. اکنون با استفاده از عملگر ضرب کرونکر، می‌توان تعداد سطرها را به هر حاصل ضربی از توان‌های اعداد اول تعمیم داد. به این ترتیب می‌توان ماتریس‌های با تعداد سطرهای دلخواه تولید کرد. نکته منفی در باره ضرب کرونکر، افزایش فاصله میان ابعاد ماتریس حاصل و کران پیش‌بینی شده برای ماتریس‌های تصادفی است. برای ماتریس $\mathbf{X}_{m_x \times n_x}$، اگر $\mathcal{B}_{\mathbf{X}} \triangleq \frac{k_x \log n_x}{m_x}$ که k_x مرتبه RIP ماتریس است. داریم:

$$\begin{aligned} \mathcal{B}_{\mathbf{C}} &= \frac{\min\{k_a, k_b\}\big(\log n_a + \log n_b\big)}{m_a m_b} \leq \frac{1}{m_b}\frac{k_a \log n_a}{m_a} + \frac{1}{m_a}\frac{k_b \log n_b}{m_b} \\ &= \frac{1}{m_b}\mathcal{B}_{\mathbf{A}} + \frac{1}{m_a}\mathcal{B}_{\mathbf{B}} \end{aligned} \tag{۵۵-۴}$$

از این رو، حتی اگر $\mathcal{B}_{\mathbf{A}}$ و $\mathcal{B}_{\mathbf{B}}$ در حد به سمت اعداد ثابت ناصفر میل کنند، $\mathcal{B}_{\mathbf{C}}$ به سمت صفر میل خواهد کرد .

۶-۴ بازسازی سریع

روش MP یکی از ساده‌ترین روش‌ها برای بازسازی بردار تنک از روی نمونه‌هاست. در اینجا نشان می‌دهیم که نمونه‌های حاصل از ماتریس‌های معرفی‌شده در این فصل، هنگامی که بدون نویز باشند توسط این روش و گونه‌های متفاوت آن قابل بازسازی هستند.

فرض کنید $\mathbf{A}_{m \times n}$ ماتریسی با ضریب همدوسی کمتر از $\frac{1}{2k-1}$ باشد که چرخش‌های دوری ستون‌های آن، مجدداً ستونی از $\mathbf{A}$ باشند. همچنین فرض کنید $\mathbf{s}_{n\times 1}$ برداری k-تنک با درایه‌های ناصفر در مکان‌های $\mathcal{S} = \{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ باشد و داشته باشیم:

$$\mathbf{y}_{m\times 1} = \mathbf{A} \cdot \mathbf{s} = \sum_{j=1}^{k} s_{i_j} \mathbf{a}_{i_j} \tag{۵۶-۴}$$

که در آن $\mathbf{a}_i$ ستون iام ماتریس $\mathbf{A}$ را نمایش می‌دهد. در روش‌های حریص، به ویژه در روش MP، تخمین بردار تنک ($\hat{\mathbf{s}}_{n\times 1}$) در ابتدا با بردار تمام صفر پایه ریزی می‌شود و در چندین تکرار به مقدار نهایی خود می‌رسد. در هر مرحله، بردار باقیمانده به صورت $\mathbf{r}_{m\times 1} = \mathbf{A}(\mathbf{s} - \hat{\mathbf{s}}) = \mathbf{y} - \mathbf{A}\hat{\mathbf{s}}$ تعریف می‌شود که به وضوح در ابتدا برابر با $\mathbf{y}$

است. در هر تکرار، ضرب داخلی بردار باقیمانده با تمام ستون‌های $\mathbf{A}$ محاسبه می‌شود و اندیس ستونی که حداکثر ضرب داخلی را ایجاد می‌کند (i_{max})، به عنوان یکی از مکان‌های ناصفر بردار $\mathbf{s}$ انتخاب می‌شود. سپس مقدار تمامی مکان‌هایی که تا این مرحله به عنوان مکان‌های ناصفر $\mathbf{s}$ انتخاب شده‌اند، بر حسب یک قاعده که در روش‌های مختلف متفاوت است، به‌روز می‌شوند و این روند تا هنگام رسیدن به شرط خاتمه ادامه می‌یابد. در اینجا نشان می‌دهیم که در هیچ یک از مراحل، در شناسایی مکان‌های ناصفر بردار $\mathbf{s}$ مرتکب اشتباه نمی‌شویم. در نتیجه اگر به‌روز رسانی مقادیر به طرز درستی انجام شود، پس از k مرحله، بردار اصلی به صورت کامل بازسازی شده است؛ در صورت به‌روز رسانی نادرست، ممکن است در چندین مرحله، موقعیت یکسانی به عنوان مکان ناصفر انتخاب شود. این موضوع را به کمک استقرا ثابت می‌کنیم: فرض کنید تا تکرار t موقعیت‌های درستی به عنوان مکان‌های ناصفر $\mathbf{s}$ انتخاب شده باشند. در نتیجه در شروع تکرار tام، مکان‌های ناصفر $\hat{\mathbf{s}}$ و $\delta = \mathbf{s} - \hat{\mathbf{s}}$ زیرمجموعه‌ای از $\mathcal{S}$ هستند. بدون کاسته شدن از کلیت مسأله فرض کنید $|\delta_{i_1}| \geq |\delta_{i_2}| \geq \cdots \geq |\delta_{i_k}|$ و سایر δ_iها صفر هستند. داریم:

$$\left|\langle \mathbf{r}, \mathbf{a}_{i_1} \rangle\right| = \left|\langle \sum_{j=1}^{k} \delta_{i_j} \mathbf{a}_{i_j}, \mathbf{a}_{i_1} \rangle\right| \geq |\delta_{i_1}| |\langle \mathbf{a}_{i_1}, \mathbf{a}_{i_1} \rangle - \sum_{j=2}^{k} |\delta_{i_j}| |\langle \mathbf{a}_{i_j}, \mathbf{a}_{i_1} \rangle| \qquad (۵۷-۴)$$

با استفاده از شرط ضریب همدوسی ماتریس $\mathbf{A}$ می‌دانیم:

$$\left|\langle \mathbf{r}, \mathbf{a}_{i_1} \rangle\right| > |\delta_{i_1}| - \frac{1}{2k-1} \sum_{j=2}^{k} |\delta_{i_j}| \geq |\delta_{i_1}| - \frac{k-1}{2k-1} |\delta_{i_1}| = \frac{k}{2k-1} |\delta_{i_1}| \qquad (۵۸-۴)$$

از طرفی اگر $l \notin \mathcal{S}$، داریم:

$$\left|\langle \mathbf{r}, \mathbf{a}_l \rangle\right| = \left|\sum_{j=1}^{k} \delta_{i_j} \langle \mathbf{a}_{i_j}, \mathbf{a}_l \rangle\right| < \frac{1}{2k-1} \sum_{j=1}^{k} |\delta_{i_j}| \leq \frac{k}{2k-1} |\delta_{i_1}| \qquad (۵۹-۴)$$

از ترکیب (۴-۵۸) و (۴-۵۹) به دست می‌آوریم:

$$\left|\langle \mathbf{r}, \mathbf{a}_l \rangle\right| \; < \; \frac{k}{2k-1} |\delta_{i_1}| \; < \; \left|\langle \mathbf{r}, \mathbf{a}_{i_1} \rangle\right| \qquad (۶۰-۴)$$

بنابراین بیشترین ضرب داخلی یا با $\mathbf{a}_{i_1}$ حاصل می‌شود یا با یکی دیگر از $\mathbf{a}_{i_j}$ها. به عبارت بهتر، بیشترین ضرب داخلی همواره برای یک عضو $\mathbf{s}_{n\times 1}$ بدست می‌آید. به این ترتیب حکم استقراء ثابت می‌شود.

همان طور که در بالا شرح داده شد، در هر تکرار، ضرب داخلی $\mathbf{r}_{m\times 1}$ با تمامی ستون‌های $\mathbf{A}$ محاسبه می‌شود. هر ضرب داخلی نیازمند m عمل ضرب و $m-1$ عمل جمع است. اکنون نشان می‌دهیم خاصیت

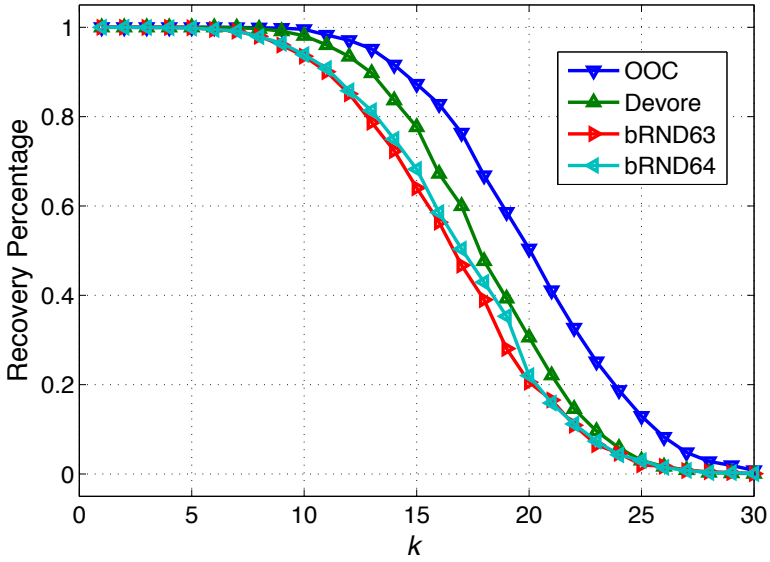
گردشی ستون‌های $\mathbf{A}$ میزان محاسبات لازم را به شدت کاهش می‌دهد. فرض کنید $\mathbf{a}$ یکی از ستون‌های ماتریس $\mathbf{A}$ و $\mathbf{a}^{(j)}$ همین ستون پس از j واحد چرخش دوری باشد. به دلیل خاصیت گردشی $\mathbf{A}$، تمامی $\mathbf{a}^{(j)}$ها ستون‌هایی از $\mathbf{A}$ هستند. در نتیجه باید $\langle \mathbf{a}^{(j)}, \mathbf{r} \rangle$ برای تمام jها محاسبه شود. فرض کنید $\{\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \ldots, \mathbf{a}^{(\mu)}\}$ گردش‌های دوری متفاوت $\mathbf{a}$ باشند (باید توجه داشت $\mu | m$). در این صورت محاسبه ضرب داخلی‌های ناشی از مجموعه مذکور و بردار $\mathbf{r}$، مستلزم μm عمل ضرب و $\mu(m-1)$ عمل جمع است. یک روش سریع برای محاسبه این ضرب داخلی‌ها استفاده از الگوریتم FFT است. نکته کلیدی در آن است که این ضرب داخلی‌ها، معادل با کانولوشن دایروی $\mathbf{r}$ و $\mathbf{a}$ هستند:

$$\langle \mathbf{r}, \mathbf{a}^{(j)} \rangle = \left(\mathbf{r} \circledast_m \mathbf{a}\right)\big|_j \qquad (۶۱-۴)$$

که نماد $\circledast_m$ کانولوشن دایروی با دوره تناوب m را نشان می‌دهد. می‌دانیم که یک روش کاراً برای محاسبه کانولوشن دایروی استفاده از ضرایب DFT است: اگر $\mathbf{r}_f$ و $\mathbf{a}_f$ به ترتیب تبدیل‌های DFT دو بردار $\mathbf{r}$ و $\mathbf{a}$ را نشان دهند، داریم:

$$IDFT\{\mathbf{r}_f \odot \mathbf{a}_f\} = \left[\left(\mathbf{r} \circledast_m \mathbf{a}\right)\big|_0, \ldots, \left(\mathbf{r} \circledast_m \mathbf{a}\right)\big|_{m-1}\right] \qquad (۶۲-۴)$$

که $\mathbf{v}_{m\times 1} \odot \mathbf{u}_{m\times 1} \triangleq [v_1 u_1, \ldots, v_m u_m]^T$. برای محاسبه این ضریب داخلی‌ها، $\mathbf{r}_f$ (در هر تکرار) تنها یک بار محاسبه می‌شود و به جای μ ضرب داخلی ذکر شده به یک عمل DFT، یک عمل IDFT و m عمل ضرب نیاز داریم. از آن‌جا که بردار $\mathbf{a}$، تنها μ گردش دوری متفاوت دارد، در حوزه DFT نیز حداکثر μ ضریب ناصفر خواهد داشت که این ضرایب مربوط به فرکانس‌هایی با فواصل یکسان هستند. در نتیجه محاسبه DFT μ–نقطه‌ای $\mathbf{a}$ کفایت می‌کند (تبدیل DFT m–نقطه‌ای بردار $\mathbf{r}$ به DFT μ–نقطه‌ای به راحتی امکان پذیر است). با استفاده از الگوریتم FFT برای عمل‌های DFT و IDFT μ–نقطه ای، به $2\mu\lceil \log_2 \mu \rceil$ عملگر ضرب و $m - \mu + \mu\lceil \log_2 \mu \rceil$ عملگر جمع نیاز داریم. مقایسه تعداد عملگرهای لازم در دو حالت فوق (استفاده از الگوریتم FFT و محاسبه مستقیم کانولوشن دوری) نشان دهنده کاهش پیچیدگی محاسباتی در عمل بازسازی است. این کاهش چشمگیر در پیچیدگی محاسباتی مدیون خاصیت گردشی ستون‌های ماتریس $\mathbf{A}$ است.

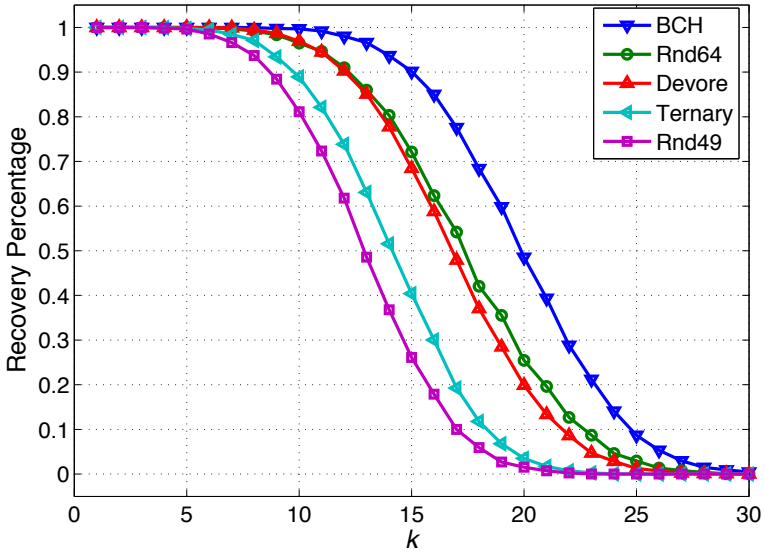


شکل ۴-۴: درصد بازسازی ($\text{SNR}_{rec.} \geq ۱۰۰dB$) در مقادیر مختلف k. ماتریس‌های حسگر با ساختار OOC و Devore ابعاد ۳۷۸ × ۶۳ و ۳۷۸ × ۶۴ دارند. همچنین دو ماتریس تصادفی دودویی با ابعاد ۳۷۸ × ۶۴ و ۳۷۸ × ۶۳ در نظر گرفته شده‌اند.

۴-۷ نتایج شبیه‌سازی

در بخش‌های پیشین، تعدادی ماتریس حسگر یقیینی معرفی کردیم. در این بخش به کمک نتایج شبیه‌سازی، عملکرد برخی از این ماتریس‌ها را بررسی می‌کنیم. از آن‌جا که ماتریس‌های مختلط، اطلاعات بیشتری در مورد سیگنال تنک ارائه می‌دهند (به کمک دو قسمت حقیقی و موهومی)، در نتایج هر شبیه‌سازی، تنها یکی از دو خانواده‌ی ماتریس‌های حقیقی و مختلط مورد بررسی قرار می‌گیرند. به عبارت دیگر، مقایسه‌ای میان عملکرد این دو دسته از ماتریس‌ها صورت نمی‌گیرد.

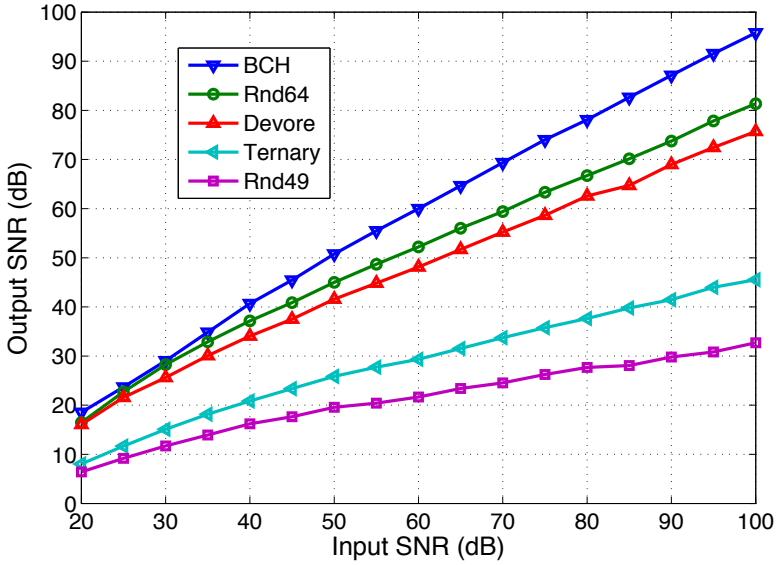
ابتدا به مقایسه عملکرد ماتریس‌های دودویی می‌پردازیم. برای این منظور از دو طرح یقینی OOC و Devore و دو ساختار تصادفی استفاده می‌کنیم. کدهای OOC معمولا با پارامتر λ کوچک طراحی می‌شوند و به همین دلیل مقایسه عادلانه بین دو ساختار OOC و Devore کمی مشکل است. برای مقایسه، از یک کد متعامد نوری (۶۳, ۹, ۲) با ۶ کلمه کد، یک ماتریس دودویی ۳۷۸ × ۶۳ با ضریب همدوسی $\frac{۲}{۹}$ ساخته‌ایم. همچنین با کمک ساختار Devore می‌توان یک ماتریس ۵۱۲ × ۶۴ با وزن ستونی ۸ و ضریب همدوسی $\frac{۱}{۴}$ ساخت که در اینجا برای مقایسه تنها ۳۷۸ ستون اول آن مورد استفاده قرار گرفته‌اند. همچنین از دو ماتریس دودویی تصادفی با



شکل ۴–۵: درصد بازسازی ($\text{SNR}_{rec.} \geq ۱۰۰dB$) در مقادیر مختلف k. ماتریس‌های حسگر با ساختار BCH، Devore و ادغامی (Ternary) ابعاد ۵۱۲ × ۶۳، ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ دارند. همچنین دو ماتریس تصادفی گوسی ابعاد ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ دارند.

ابعاد ۳۷۸ × ۶۳ و ۳۷۸ × ۶۴ بهره جسته‌ایم به نحوی که درایه‌ها به صورت مستقل و به ترتیب با احتمال $\frac{۹}{۶۴}$ و $\frac{۸}{۶۴}$ مقدار یک اختیار می‌کنند. شکل ۴–۴ درصد بازسازی کامل ($\text{SNR}_{rec} \geq ۱۰۰dB$) را هنگام پیاده‌سازی روش OMP و استفاده از ماتریس‌های فوق برای مرتبه‌های تنک‌بودن (k) متفاوت نشان می‌دهد. در این شبیه‌سازی، نمونه‌ها بدون نویز هستند و نمودارها از میانگین گیری ۵۰۰۰ شبیه‌سازی متفاوت حاصل شده‌اند. همانگونه که از مقادیر ضرایب همدوسی انتظار می‌رود ساختار OOC بهترین عملکرد را دارد.

برای بررسی ماتریس‌های حقیقی، از ماتریس‌های زیر استفاده کرده‌ایم: ماتریس دودویی ۵۱۲ × ۶۴ با ساختار Devore، ماتریس دو قطبی ۵۱۲ × ۶۳ با ساختار BCH، ادغام ماتریس دودویی ۳۴۳ × ۴۹ با ساختار Devore و ماتریس دو قطبی ۸ × ۷ با ساختار BCH. پس از ادغام، ماتریسی با ابعاد ۲۷۴۴ × ۴۹ حاصل می‌شود که ما تنها ۵۱۲ ستون اول آن را در نظر می‌گیریم. ضریب همدوسی این ماتریس‌ها به ترتیب $\frac{۱}{۴}$، $\frac{۱}{۷}$ و $\frac{۱}{۴}$ است. شکل ۴–۵ درصد موفقیت ($\text{SNR}_{rec.} \geq ۱۰۰dB$) روش OMP در بازسازی سیگنال تنک از روی نمونه‌های حاصل شده به کمک این ماتریس‌ها را نمایش می‌دهد. به منظور مقایسه جامع تر، ماتریس‌های تصادفی با توزیع گوسی با اندازه‌های ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ نیز استفاده شده‌اند و نتایج این بازسازی‌ها در ۵۰۰۰ شبیه‌سازی

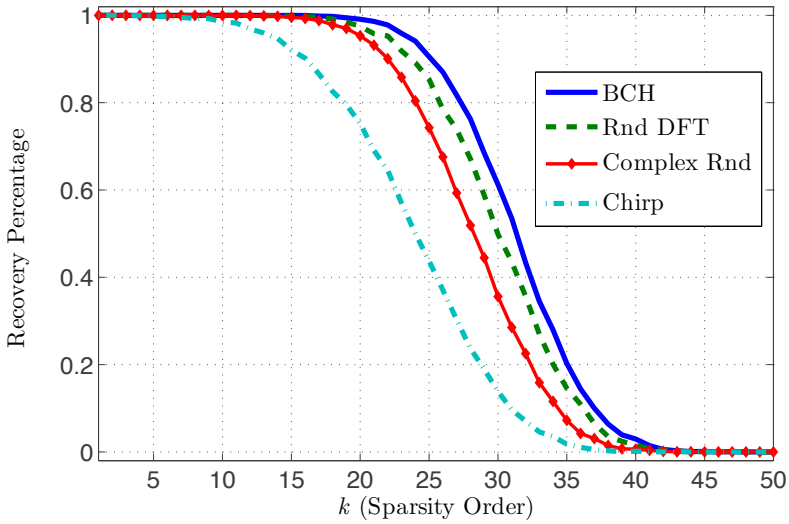


شکل ۴–۶: SNR سیگنال‌های ۱۵–تنک بازسازی شده هنگامی که نمونه‌های فشرده تحت تاثیر نویز جمعی با توان‌های متفاوتی قرار گیرند. ماتریس‌های حسگر با ساختار BCH، Devore و ادغامی (Ternary) ابعاد ۵۱۲ × ۶۳، ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ دارند. همچنین دو ماتریس تصادفی گوسی ابعاد ۵۱۲ × ۶۴ و ۵۱۲ × ۴۹ دارند.

متفاوت متوسط گیری شده است. در شکل ۴–۶ عملکرد همین ماتریس‌ها در شرایط نویزی و با میزان نویزهای متفاوت و هنگامی که سیگنال ورودی ۱۵–تنک است، نشان داده شده‌اند. در هر دو شکل، ماتریس‌های دو قطبی با ساختار BCH نسبت به ماتریس‌های تصادفی متناظر و ماتریس دودویی Devore عملکرد بهتری دارند.

برای ماتریس‌های مختلط، از ماتریس‌های BCH با ابعاد $p^۶ \times (p^۴ - ۱)$ به ازای $p = ۳,۵$ استفاده کرده‌ایم. همچنین، ماتریس‌های بر مبنای توابع Chirp، ماتریس‌های تصادفی با توزیع گوسی و زیر ماتریس‌های تصادفی از ماتریس DFT (سطرهای تصادفی) با اندازه‌های مشابه پیاده‌سازی شده‌اند. شکل ۴–۷ و ۴–۸ نتایج در حالت $p = ۳$ و شکل ۴–۹ نتایج در حالت $p = ۵$ را نشان می‌دهند. در تمام این شکل‌ها، عملکرد ماتریس BCH در حد زیر ماتریس‌های تصادفی از ماتریس DFT است که از عملکرد سایر روش‌ها بهتر است.

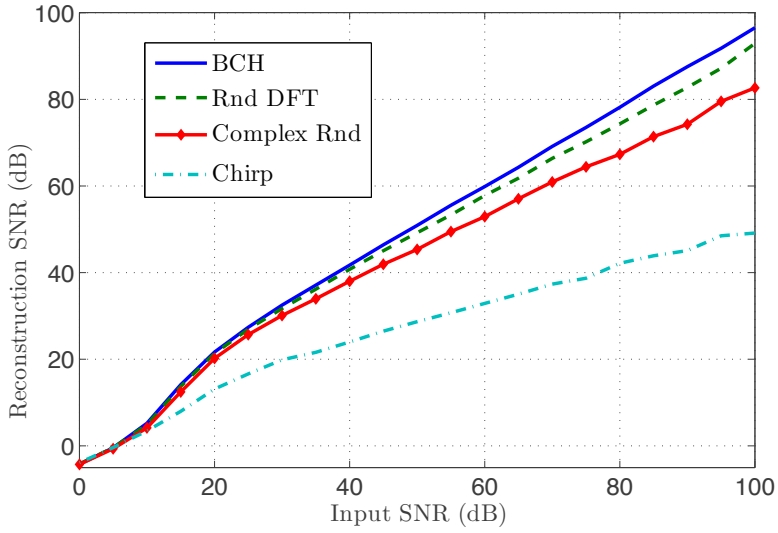
در شکل ۴–۱۰ عملکرد ماتریس‌های ادغامی را مورد مطالعه قرار می‌دهیم. برای این منظور، ماتریس دودویی ۵۱۲ × ۶۴ با ساختار Devore و وزن ۸ را با ماتریس ۹ × ۸ از نوع BCH ۳–سمبلی ادغام می‌کنیم تا یک ماتریس ۴۶۰۸ × ۶۴ با ضریب همدوسی $\frac{۱}{۴}$ بدست آید. به عنوان مثالی از ضرب کرونکر ماتریس‌ها، یک ماتریس دودویی ۲۷ × ۹ با ساختار Devore و وزن ستونی ۳ را با یک ماتریس دو قطبی ۶۴ × ۷ با ساختار BCH ادغام



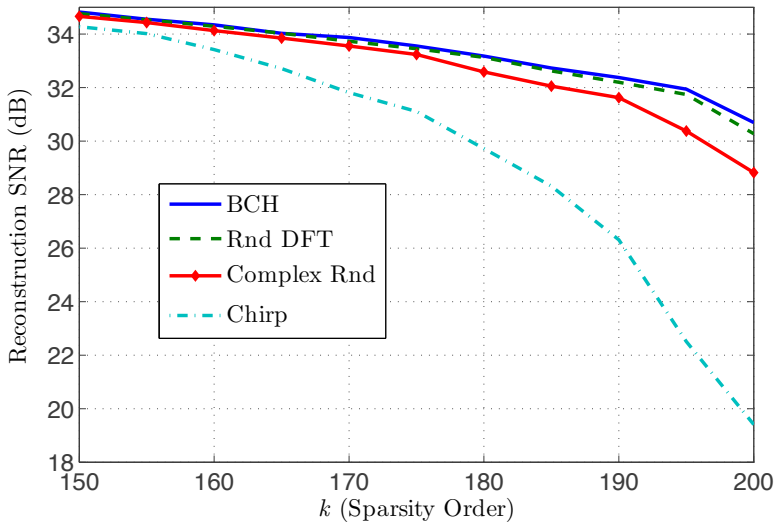
شکل ۴-۷: درصد بازسازی کامل ($\text{SNR}_{rec.} \geq ۱۰۰dB$) هنگامی که نمونه‌ها بدون نویز هستند. ماتریس BCH بر مبنای کدهای $p = ۳$ سمبلی ساخته شده است و ابعاد تمامی ماتریس‌ها ۷۲۹ × ۸۰ است.

کرده‌ایم به طوری که یک ماتریس ۱۷۲۸ × ۶۳ با ضرب همدوسی $\frac{۵}{۷}$ حاصل شود. ماتریس‌های تصادفی در این شکل با ابعاد ۴۶۰۸ × ۶۴ هستند اما ماتریس بر پایه توابع Chirp ۴۶۰۸ × ۷۵ است. نتایج شبیه‌سازی حاکی از اختلاف شدید بین دو روش ادغام ماتریس است. عملکرد روش ادغام با ماتریس‌های دودویی کمابیش در حد عملکرد ماتریس‌های تصادفی است در حالی که عملکرد ماتریس حاصل از ضرب کرونکر بسیار ناامیدکننده است.

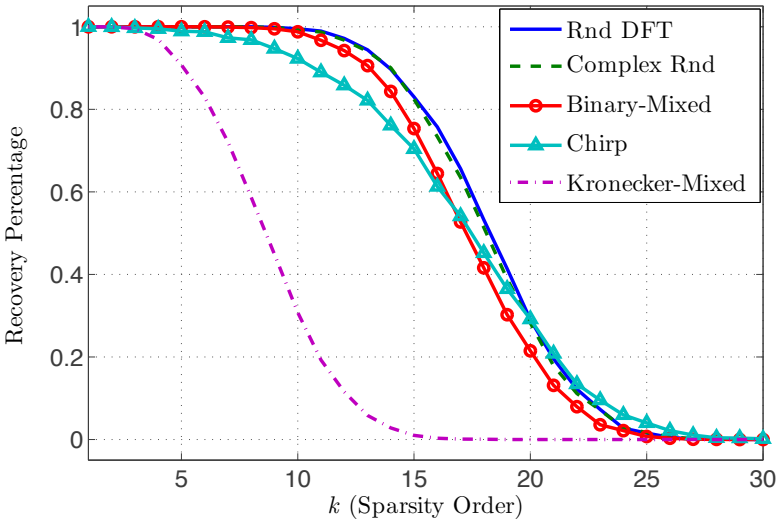
در انتها برای مقایسه پیچیدگی محاسباتی روش OMP با و بدون استفاده از الگوریتم FFT، شکل ۴-۱۱ را رسم کرده‌ایم. در این شکل زمان بازسازی یک سیگنال ۱۵۶۲۵ × ۱ توسط نمونه‌های ۶۲۴ بعدی آن نشان داده شده است. برای حالت $k = ۴۵$ زمان لازم برای OMP ساده بیش از ۱۲ برابر زمان لازم برای حالتی است که از FFT استفاده کنیم.



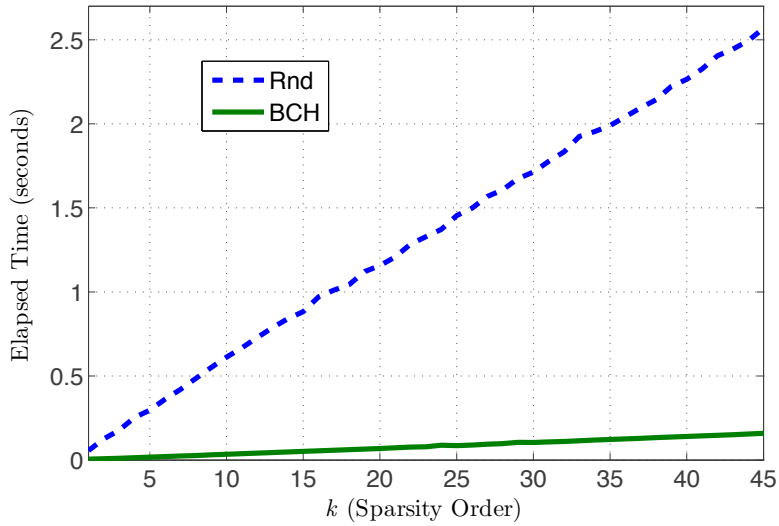
شکل ۴-۸: SNR سیگنال ۲۵-تنک بازسازی شده برای توان‌های متفاوت نویز جمعی. ماتریس BCH بر مبنای کدهای $p = ۳$ سمبلی ساخته شده است و ابعاد تمامی ماتریس‌ها ۷۲۹ × ۸۰ است.



شکل ۴-۹: SNR سیگنال ۲۵-تنک بازسازی شده برای توان‌های متفاوت نویز جمعی. ماتریس BCH بر مبنای کدهای $p = ۵$ سمبلی ساخته شده است و ابعاد تمامی ماتریس‌ها ۱۵۶۲۵ × ۶۲۴ است.



شکل ۴-۱۰: درصد بازسازی کامل ($\text{SNR}_{rec.} \geq ۱۰۰dB$) هنگامی که نمونه‌ها بدون نویز هستند. ابعاد ماتریس‌های بر مبنای ادغام دودویی، ادغام کرونکر، توابع Chirp، تصادفی گوسی و سطرهای تصادفی ماتریس DFT به ترتیب عبارتند از ۴۶۰۸ × ۶۴، ۱۷۲۸ × ۶۳، ۴۶۰۸ × ۷۵، ۴۶۰۸ × ۶۴ و ۴۶۰۸ × ۶۴.



شکل ۴-۱۱: مقایسه زمان لازم برای بازسازی یک بردار تنک ۱ × ۱۵۶۲۵ از نمونه‌های فشرده ۱ × ۶۲۴ توسط ماتریس تصادفی (OMP ساده) و ماتریس‌های BCH با $p = ۵$ سمبلی (OMP تسریع شده).

فصل ۵

نمونه‌برداری غیرتصادفی غیرخطی

۵-۱ مقدمه

در فصل‌های پیشین دیدیم که ماتریس‌های تصادفی، مستقل از حوزه تنک‌بودن سیگنال (ماتریس $\mathbf{\Psi}$)، قادر به انجام عمل نمونه‌برداری با احتمال بازسازی مطلوب (به اندازه کافی بزرگ) هستند. از سوی دیگر، در صورت معلوم بودن حوزه تنک‌بودن، روش‌هایی برای طراحی ماتریس حسگر یقینی معرفی کردیم که بازسازی کامل را تضمین می‌کنند. برای این منظور مساله را در دو حالت معلوم و یا نامعلوم بودن حوزه تنک‌بودن بررسی می‌کنیم. در حالت حوزه معلوم، بررسی می‌کنیم که آیا با غیرخطی کردن نحوه نمونه‌برداری می‌توان در نرخ فشرده‌سازی و یا پیچیدگی محاسباتی در بازسازی به بهبودی دست یافت.

قسمتی از مطالب این فصل بر گرفته از [۵] هستند.

۵-۲ نمونه‌برداری غیرتصادفی برای حوزه تنک‌بودن نامعلوم

در این بخش به دنبال بررسی مساله نمونه‌برداری یقینی، هنگام ناشناخته بودن حوزه تنک‌بودن سیگنال هستیم؛ به بیان دیگر، اگر $\mathbf{\Psi}$ در هنگام نمونه‌برداری معلوم نباشد، آیا می‌توان از یک روش نمونه‌برداری ثابت و غیرتصادفی استفاده کرد بدون این که قابلیت بازسازی از بین برود؟ در این‌جا، نمونه‌برداری لزوماً به روش‌های خطی مثل استفاده از یک ماتریس $\mathbf{\Phi}$ محدود نیست. یعنی در حالت کلی، در پی یافتن روش نمونه‌برداریی هستیم که مستقل از حوزه تنک‌بودن سیگنال، بتوان آن را بازسازی کرد. برای بررسی، ابتدا از روش خطی ضرب ماتریس $\mathbf{\Phi}$ (مشابه فصل قبل) آغاز می‌کنیم؛ اگر $\mathbf{\Phi}_{m\times n}$ ماتریس ثابتی با $m \geq n$ باشد، تعداد نمونه‌ها بیشتر یا مساوی نمونه‌های اولیه

سیگنال است که مسلماً متناقض با تمام اهداف نمونه‌برداری فشرده است. در حالت $m < n$ نیز فضای پوچ $\mathbf{\Phi}$ بعد $n - m > ۰$ خواهد داشت. در نتیجه:

$$\exists \mathbf{v} \in \mathbb{R}^n,\ \|\mathbf{v}\|_{\ell_۲} = ۱ : \quad \mathbf{\Phi} \cdot \mathbf{v} = ۰ \tag{۱–۵}$$

فرض کنید $V^\perp$ فضای حاصل از تمام بردارهای عمود بر $\mathbf{v}$ در $\mathbb{R}^n$ باشد؛ در نتیجه بعد $V^\perp$ برابر با $n-۱$ است و می‌توان بردارهای متعامد یکه $\{\mathbf{u}_۱, \ldots, \mathbf{u}_{n-۱}\}$ را به عنوان پایه برای $V^\perp$ اختیار کرد. بنابراین مجموعه بردارهای $\{\mathbf{v}, \mathbf{u}_۱, \ldots, \mathbf{u}_{n-۱}\}$ برای $\mathbb{R}^n$، تشکیل یک پایه متعامد یکه می‌دهند. اگر چنین تعریفی داشته باشیم:

$$\mathbf{\Psi} = [\mathbf{v}\ \mathbf{u}_۱\ \ \ldots\ \ \mathbf{u}_{n-۱}] \Rightarrow \mathbf{v} = \mathbf{\Psi} \cdot \begin{bmatrix} ۱ \\ ۰ \\ \vdots \\ ۰ \end{bmatrix} \tag{۲–۵}$$

در نتیجه، $\mathbf{v}$ در حوزه متعامد یکه $\mathbf{\Psi}$ به صورت ۱–تنک بیان می‌شود اما $\mathbf{\Phi} \cdot \mathbf{v} = ۰$؛ بنابراین بردار ۱–تنکی وجود دارد که توسط نمونه‌های تولید شده با $\mathbf{\Phi}$ قابل بازسازی نیست. در نتیجه، در هر نوع نمونه‌برداری خطی، بردارهای تنکی وجود دارند که قابل بازسازی نیستند. اکنون به نمونه‌برداری‌های غیرخطی می‌پردازیم؛ به عبارت دیگر، درپی توابع نمونه‌برداری $f_۱, \ldots, f_n$ هستیم که از روی نمونه‌های تولید شده آن‌ها بتوان بردار تنک ورودی را بدست آورد:

$$\begin{cases} y_۱ = f_۱(\mathbf{x}) = f_۱(x_۱, \ldots, x_n) \\ \qquad\qquad \vdots \\ y_m = f_m(\mathbf{x}) = f_m(x_۱, \ldots, x_n) \end{cases} \tag{۳–۵}$$

نکته مهمی که باید به آن اشاره کرد این است که $\mathbb{R}^n$ و $\mathbb{R}$ از نظر Cardinality با هم برابرند؛ یعنی تابع یک به یک و پوشای $g : \mathbb{R}^n \to \mathbb{R}$ وجود دارد که تمامی بردارهای n تایی را به طور یک به یک با اعداد حقیقی، متناظر می‌کند. اگر از این تابع برای نمونه‌برداری استفاده شود، تمام n نمونه‌ی اولیه را حتی بدون نیاز به شرط تنک‌بودن، در یک نمونه نهایی، فشرده کرده‌ایم. با وجود قابلیت فشرده‌سازی حیرت‌انگیز این تابع، مشکلی اساسی وجود دارد: چنانچه نمونه نهایی چندی شود، تمام اطلاعات (تمام n نمونه‌ی اولیه) از بین می‌روند! بدیهی است که این قابلیت فشرده‌سازی بدون چندی‌کردن هیچ ارزش عملی ندارد. به بیان ساده‌تر، روشی مورد نظر است که پس از چندی‌کردن نمونه‌ها، هنوز قادر به بازسازی سیگنال اولیه (با احتساب خطایی محدود) باشد. در نتیجه، حتماً باید از شرط تنک‌بودن استفاده شود؛ یعنی بردار $\mathbf{x}$ تنها برداری از $\mathbb{R}^n$ نباشد که نمونه‌های $\mathbf{y}_{m\times ۱}$

را تولید می‌کند بلکه تنها بردار تنک با این خاصیت باشد. نشان خواهیم داد که اگر هدف بازسازی سیگنال‌های k-تنک برای $k \geq ۲$ باشد، حتی توابع غیرخطی (با قابلیت چندی‌کردن) نیز کارآمد نخواهند بود؛ اما برای حالت $k = ۱$ چنین توابعی قابل ارائه است. فرض کنید $k \geq ۲$ و $(y_۱, \ldots, y_m)$ نمونه‌های بدست آمده از بردار $\mathbf{x}_n \times ۱$ با توابع نمونه‌برداری $(f_۱, \ldots, f_m)$ باشند، تعریف کنید:

$$S = \{\mathbf{z} \in \mathbb{R}^n \mid \forall ۱ \leq i \leq m : \ f_i(\mathbf{z}) = y_i\} \tag{۴-۵}$$

به بیان دیگر، مجموعه S شامل تمام بردارهایی است که نمونه‌های تولید شده آن‌ها با نمونه‌های تولید شده از $\mathbf{x}_{n\times ۱}$ برابرند و در نتیجه ابهام ما در بازسازی به انتخاب یکی از اعضای S محدود می‌شود. از آنجا که $\mathbf{x}_{n\times ۱} \in S$ پس $S \neq \emptyset$. از طرف دیگر مشابه استدلالی که در قبل به آن اشاره شد، $\mathbf{x}_{n\times ۱}$ نمی‌تواند تنها عضو S باشد زیرا در این صورت بدون توجه به شرط تنک‌بودن، سیگنال $\mathbf{x}_{n\times ۱}$ بازسازی می‌شود:

$$\exists \mathbf{v}_۰ \in S, \ \mathbf{v}_۰ \neq \mathbf{x}_{n\times ۱} \tag{۵-۵}$$

تعریف کنید:

$$\mathbf{v}_۱ = \mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{v}_۰ \rangle}{\langle \mathbf{v}_۰, \mathbf{v}_۰ \rangle} \mathbf{v}_۰ \tag{۶-۵}$$

که منظور از $\langle \mathbf{a}, \mathbf{b} \rangle$ ضرب داخلی دو بردار $\mathbf{a}$،$\mathbf{b}$ است. فرض کنید $\tilde{\mathbf{v}}_۰$،$\tilde{\mathbf{v}}_۱$ بردارهای یکه در راستاهای $\mathbf{v}_۰$،$\mathbf{v}_۱$ باشند. همچنین فرض کنید $V^\perp$ فضای ساخته شده از تمام بردارهای عمود بر $\tilde{\mathbf{v}}_۰$،$\tilde{\mathbf{v}}_۱$ و $\{\mathbf{u}_۱, \ldots, \mathbf{u}_{n-۲}\}$ پایه‌ای متعامد یکه برای این فضا باشد:

$$\begin{aligned} \langle \mathbf{v}_۰, \mathbf{v}_۱ \rangle &= \langle \mathbf{x}, \mathbf{v}_۰ \rangle - \frac{\langle \mathbf{x}, \mathbf{v}_۰ \rangle}{\langle \mathbf{v}_۰, \mathbf{v}_۰ \rangle} \langle \mathbf{v}_۰, \mathbf{v}_۰ \rangle = ۰ \\ &\Rightarrow \quad \mathbf{v}_۰ \perp \mathbf{v}_۱ \Rightarrow \tilde{\mathbf{v}}_۰ \perp \tilde{\mathbf{v}}_۱ \end{aligned} \tag{۷-۵}$$

بنابراین، $\{\tilde{\mathbf{v}}_۰, \tilde{\mathbf{v}}_۱, \mathbf{u}_۱, \ldots, \mathbf{u}_{n-۲}\}$ تشکیل یک پایه متعامد یکه برای $\mathbb{R}^n$ می‌دهند. اگر ماتریس $\mathbf{\Psi}$ را از کنار هم قراردادن این بردارها بوجود آوریم ($\mathbf{\Psi} = [\tilde{\mathbf{v}}_۰ \ \tilde{\mathbf{v}}_۱ \ \mathbf{u}_۱ \ \ldots \ \mathbf{u}_{n-۲}]$)، ماتریس $\mathbf{\Psi}$ متعامد یکه خواهد بود و

بردارهای $\mathbf{x}$ و $\mathbf{v}_\circ$ در این حوزه، نمایش تنک دارند:

$$\mathbf{v}_\circ \quad = \quad \|\mathbf{v}_\circ\|_{\ell_2}\tilde{\mathbf{v}}_\circ \Rightarrow \boldsymbol{\Psi}\cdot\begin{bmatrix}\|\mathbf{v}_\circ\|_{\ell_2}\\ \circ\\ \vdots\\ \circ\end{bmatrix} = \mathbf{v}_\circ$$

$$\mathbf{x} = \mathbf{v}_1 + \frac{\langle \mathbf{x},\mathbf{v}_\circ\rangle}{\langle \mathbf{v}_\circ,\mathbf{v}_\circ\rangle}\mathbf{v}_\circ \quad = \quad \|\mathbf{v}_1\|_{\ell_2}\tilde{\mathbf{v}}_1 + \langle \mathbf{x},\mathbf{v}_\circ\rangle\tilde{\mathbf{v}}_\circ \Rightarrow \boldsymbol{\Psi}\cdot\begin{bmatrix}\langle \mathbf{x},\mathbf{v}_\circ\rangle\\ \|\mathbf{v}_1\|_{\ell_2}\\ \circ\\ \vdots\\ \circ\end{bmatrix} = \mathbf{x} \qquad (۸–۵)$$

یعنی $\mathbf{x}$ و $\mathbf{v}_\circ$ در این حوزه به ترتیب ۲–تنک و ۱–تنک هستند و از آنجا که $k \geq 2$ است، هر دو این پاسخ‌ها قابل قبول‌اند؛ بنابراین مسئله جواب یکتا ندارد. اگر $k = 1$ باشد، از آنجا که یکی از پاسخ‌های فوق ۲–تنک است، قابل قبول نیست و در نتیجه حالت $k = 1$ باید جداگانه بررسی شود. درواقع، برای $k = 1$ چنین توابعی وجود دارند که قبل از معرفی این توابع، ابتدا یک شرط لازم و کافی بیان می‌شود:

لم ۵–۱ مجموعه $S \subset \mathbb{R}^n$ قابلیت بازسازی یکتا برای سیگنال‌های ۱–تنک را نسبت به حوزه تنک‌بودن دلخواه فراهم می‌کند اگر و فقط اگر :

$$\forall \mathbf{a} \neq \mathbf{b} \in S: \quad \circ < |\langle \mathbf{a},\mathbf{b}\rangle| < \|\mathbf{a}\|_{\ell_2}\|\mathbf{b}\|_{\ell_2} \qquad (۹–۵)$$

قبل از اثبات لم فوق، باید به دو نکته اشاره کرد: اول این که قدر مطلق هر عدد حقیقی، نامنفی است؛ پس تنها نکته نابدیهی در نامساوی سمت چپ، عدم تساوی با صفر است. به عبارت بهتر، برقراری نامساوی سمت چپ ایجاب می‌کند که هیچ دو عضوی از S بر یکدیگر عمود نباشند؛ نکته دوم این است که با استفاده از نامساوی کوشی، چنین داریم:

$$\forall \mathbf{a} \neq \mathbf{b} \in \mathbb{R}^n: \quad |\langle \mathbf{a},\mathbf{b}\rangle| \leq \|\mathbf{a}\|_{\ell_2}\|\mathbf{b}\|_{\ell_2} \qquad (۱۰–۵)$$

و تساوی زمانی رخ می‌دهد که $\mathbf{a}$ و $\mathbf{b}$ در یک راستا باشند:

$$|\langle \mathbf{a},\mathbf{b}\rangle| = \|\mathbf{a}\|_{\ell_2}\|\mathbf{b}\|_{\ell_2} \quad \Longleftrightarrow \quad \exists c \in \mathbb{R}: \mathbf{a} = c\mathbf{b} \qquad (۱۱–۵)$$

بنابراین مجدداً، تنها نکته غیربدیهی در نامساوی سمت راست (۵–۹) ، عدم تساوی است؛ یعنی هیچ دو عضوی از S همراستا نیستند. اکنون نشان می‌دهیم اگر S شرایط لم فوق را ارضا کند، برای هر $\boldsymbol{\Psi}$ متعامد یکه

دلخواه، حداکثر یک عضو از S در حوزه $\mathbf{\Psi}$ نمایش ۱–تنک دارد. فرض کنید $\mathbf{a},\mathbf{b}\in S$ و هردوی $\mathbf{a},\mathbf{b}$ نمایش ۱–تنک در حوزه یک $\mathbf{\Psi}$ متعامد یکه داشته باشند؛ بنابراین هرکدام از $\mathbf{a}$،$\mathbf{b}$ ضریبی از یکی از ستون‌های $\mathbf{\Psi}$ خواهند بود. اگر هر دو، ضریبی از یک ستون خاص در $\mathbf{\Psi}$ باشند، چون راستای هردو با راستای این ستون از $\mathbf{\Psi}$ یکسان است، $\mathbf{a}$،$\mathbf{b}$ همراستا خواهند بود که مغایر با شرط لم است. در صورتی هم که ستون‌های مربوط به $\mathbf{a},\mathbf{b}$ در $\mathbf{\Psi}$ متمایز باشند، به دلیل ساختار متعامد یکه $\mathbf{\Psi}$، این دو ستون بر هم عمودند و در نتیجه $\mathbf{a},\mathbf{b}$ متعامد خواهند بود که مشابهاً متناقض با شرط لم است. پس چنین $\mathbf{a},\mathbf{b}$ای نمی‌توانند موجود باشند؛ در حوزه $\mathbf{\Psi}$ حداکثر یک عضو از S نمایش ۱–تنک دارد. اثبات عکس مطلب فوق (لازم بودن شرایط) نیز به صورت مشابهی قابل بیان است. حال باید توابعی معرفی کنیم که کلاس‌های هم‌ارزی خروجی‌های آنها مجموعه‌هایی با شرایط ذکر شده در لم باشد. نشان می‌دهیم که دو تابع نمونه‌برداری زیر چنین شرایطی را ایجاد می‌کنند:

$$\begin{cases} f_1(\mathbf{x}) &= \sum_{i=1}^{n} 3^{i-1}\operatorname{sign}(x_i) \\ f_2(\mathbf{x}) &= \|\mathbf{x}_{n\times 1}\|_{\ell_1} = \sum_{i=1}^{n}|x_i| \end{cases} \qquad (۱۲–۵)$$

که در آن:

$$\operatorname{sign}(a) = \begin{cases} 1 & a>0 \\ 0 & a=0 \\ -1 & a<0 \end{cases} \qquad (۱۳–۵)$$

در واقع تابع f_1 تنها بیانگر علامت درایه‌های بردار $\mathbf{x}_{n\times 1}$ است (به تعبیری بسط مبنای ۳ تولید شده است). می‌توان به جای تابع f_1 از n تا تابع مجزا برای علامت هر درایه استفاده کرد. از آنجا که بازه تغییرات f_1 بین $-\frac{3^n-1}{2}, \frac{3^n-1}{2}$ محدود است، برای چندی‌کردن خروجی این تابع، حداکثر $\lceil n\log_2 3\rceil$ بیت مورد نیاز است. قبل از بررسی اثر چندی‌کردن خروجی تابع f_1 ، ابتدا نشان می‌دهیم که شرایط لم صادق است. اگر $f_1(\mathbf{x})=0$ باشد، تنها حالت قابل قبول $\mathbf{x}_{n\times 1}=0$ است (با توجه به علامت درایه‌ها). بنابراین، در چنین حالتی حتی بدون نیاز به خروجی تابع f_2، ورودی نمونه‌برداری به طور یکتا بازسازی می‌شود. اگر برای دو بردار $\mathbf{a}$،$\mathbf{b}$ چنین داشته باشیم: $f_1(\mathbf{a})=f_2(\mathbf{b})\neq 0$، در این صورت علامت درایه‌های متناظر $\mathbf{a}$،$\mathbf{b}$ مشابه است و در بین درایه‌های هرکدام، حداقل یک عدد ناصفر وجود دارد :

$$\begin{cases} \mathbf{a} &= [a_1 \ \dots \ a_n]^T \\ \mathbf{b} &= [b_1 \ \dots \ b_n]^T \end{cases} \qquad f_1(\mathbf{a})=f_1(\mathbf{b}) \Rightarrow \forall 1\le i\le n:\ \operatorname{sign}(a_i)=\operatorname{sign}(b_i)$$

$$\Rightarrow \quad a_ib_i\ge 0 \Rightarrow \sum_{i=1}^{n} a_ib_i \ge 0 \Rightarrow \langle \mathbf{a},\mathbf{b}\rangle \ge 0 \qquad (۱۴–۵)$$

و تساوی هم اتفاق نمی‌افتد زیرا دست‌کم برای یک مقدار i حاصل، a_ib_i ناصفر خواهد بود:

$$\langle \mathbf{a}, \mathbf{b} \rangle > ۰ \tag{۱۵-۵}$$

به بیان دیگر، اگر $f_۱(\mathbf{a}) = f_۱(\mathbf{b})$، دو بردار $\mathbf{a}$ و $\mathbf{b}$ نمی‌توانند عمود باشند (به جز در حالت صفر). بنابراین از بین راستاهای ۱-تنکی که ماتریس متعامد یکه دلخواه $\mathbf{\Psi}$ ایجاد می‌کند، حداکثر یکی می‌تواند نمونه حاصل از $f_۱$ را تولید کند. تنها نکته باقی مانده این است که نمونه حاصل از $f_۲$ ایجاب می‌کند که در هر راستای دلخواه، حداکثر یک بردار، نمونه‌های حاصل از $f_۱$ و $f_۲$ را ایجاد کند. اگر $\mathbf{a}$ و $\mathbf{b}$ همراستا باشند، چنین است:

$$\exists c \in \mathbb{R} : \ \mathbf{a} = c\mathbf{b} \Rightarrow f_۲(\mathbf{a}) = \|\mathbf{a}\|_{\ell_۱} = |c| \ \|\mathbf{b}\|_{\ell_۱} = |c| f_۲(\mathbf{b}) \tag{۱۶-۵}$$

پس اگر $f_۲(\mathbf{a}) = f_۲(\mathbf{b})$، باید $|c| = ۱$ باشد. حالت $c = ۱$ حالت مورد نظر است و تناقضی ایجاد نمی‌کند اما حالت $c = -۱$ قابل قبول نیست:

$$c = -۱ \Rightarrow \mathbf{a} = -\mathbf{b} \Rightarrow \langle \mathbf{a}, \mathbf{b} \rangle = -\|\mathbf{b}\|_{\ell_۲}^۲ \tag{۱۷-۵}$$

که با شرط تولید شده از $f_۱(\mathbf{a}) = f_۱(\mathbf{b})$ در تضاد است. بنابراین ثابت شد که توسط نمونه‌برداری به کمک دو تابع $f_۱$ و $f_۲$ معرفی شده، مستقل از انتخاب حوزه تنک‌بودن (ماتریس $\mathbf{\Psi}$)، حداکثر یک بردار ۱-تنک بازسازی می‌شود. همان‌طور که مشاهده شد، راستای بردار ۱-تنک توسط نمونه حاصل از $f_۱$ تعیین می‌شود و در نتیجه با چندی‌کردن خروجی تابع $f_۲$ ، راستای سیگنال بازسازی شده تغییر نمی‌کند! تنها تفاوتی که بوجود می‌آید در اندازه بردار بازسازی شده است. به سهولت می‌توان مشاهده کرد که خطای تولید شده در اندازه بردار، به طور خطی متناسب با خطای ناشی از چندی‌کردن نمونه حاصل از $f_۲$ است. پس، خطاهای کوچک ورودی، به خطاهای کوچک خروجی تبدیل می‌شوند. اگر به جای بردارهای $\mathbb{R}^n$ بردارهای $\mathbb{C}^n$ مورد نظر باشند، به جای دو تابع فوق باید از سه تابع زیر استفاده کرد:

$$\begin{cases} f_۱(\mathbf{x}) &= \sum_{i=۱}^{n} ۳^{۲(i-۱)} \mathrm{sign}\big(\Re x_i\big) + ۳^{۲i-۱} \mathrm{sign}\big(\Im x_i\big) \\ f_۲(\mathbf{x}) &= \sum_{i=۱}^{n} |x_i| \\ f_۳(\mathbf{x}) &= \sum_{x_i \neq ۰} \frac{\mathrm{msign}(\Re x_i) \cdot \Im x_i}{|\Re x_i| + |\Im x_i|} \end{cases} \tag{۱۸-۵}$$

که در آن:

$$\mathrm{msign}(a) = \begin{cases} ۱ & a \geq ۰ \\ -۱ & a < ۰ \end{cases} \tag{۱۹-۵}$$

دو تابع اول تقریباً مشابه توابع قبلی هستند و نقش مشابهی ایفا می‌کنند؛ تابع سوم برای بازیابی فاز است. برای پرهیز از طولانی شدن متن، حالت $\mathbb{C}^n$ را اثبات نمی‌کنیم.

۵-۳ نمونه‌برداری غیرتصادفی برای حوزه تنک‌بودن معلوم

از ابتدایی‌ترین مفروضات نمونه‌برداری فشرده، خطی بودن نحوه‌ی نمونه‌برداری است. خطی بودن، مزایای بسیاری از جمله قابلیت تحلیل ساده و تقویت نکردن نویز جمعی را در بر دارد، اما آیا می‌توان با ورود به روش‌های نمونه‌برداری غیرخطی مزیتی بدست آورد؟ در این بخش نشان می‌دهیم که با استفاده از روش‌های غیرخطی می‌توان تعداد نمونه‌های لازم برای بازسازی در حالت بدون نویز را تا $۲k-۱$ پایین آورد که بسیار پایین‌تر از نرخ لازم در روش‌های خطی ذکر شده است. همچنین، نحوه بازسازی با پیچیدگی بسیار کمتری صورت می‌گیرد و تضمین بازسازی سیگنال اولیه به صورت قطعی است و نه به صورت احتمالاتی. در مقابل، روش‌های غیرخطی نسبت به نویز جمعی پایداری کمتری دارند و همچنین پیچیدگی نمونه‌برداری را افزایش می‌دهند.

فرض کنید $\mathbf{x}_{n\times ۱}$ یک بردار k-تنک باشد. می‌خواهیم با نمونه‌برداری غیرخطی از $\mathbf{x}_{n\times ۱}$ با تعداد نمونه کم $(۲k-۱)$ آن را بازسازی کنیم (نتایج را می‌توان به حالتی که حوزه تنک‌بودن $\mathbf{\Psi}$ مطرح است نیز تعمیم داد). فرض کنید $a_۱,\ldots,a_n$ اعداد حقیقی دلخواه باشند؛ نمونه‌های غیرخطی را به صورت زیر تعریف می‌کنیم:

$$\begin{cases} y_۱ &= a_۱x_۱ + a_۲x_۲ + \cdots + a_nx_n \\ y_۲ &= a_۱x_۱^۲ + a_۲x_۲^۲ + \cdots + a_nx_n^۲ \\ &\vdots \\ y_{۲k-۱} &= a_۱x_۱^{۲k-۱} + a_۲x_۲^{۲k-۱} + \cdots + a_nx_n^{۲k-۱} \end{cases} \qquad (۲۰\text{-}۵)$$

(با وجود شباهت ظاهری به کدینگ RS[۱] و یا ماتریس‌های واندرموند، دقت کنید که خود مجهولات به توان می‌رسند و نه ضرایب؛ بنابراین، نمونه‌برداری غیرخطی است). نشان می دهیم که هر حالت k-تنک در بین بردارهای $\mathbf{x}_{n\times ۱}$ را می‌توان از روی نمونه‌های $\{y_۱,\ldots,y_{۲k-۱}\}$ استخراج کرد. ابتدا دقت کنید که در روابط فوق حداکثر k تا از جمله‌های x_i ناصفرهستند، مثلاً $\{x_{i_۱},\ldots,x_{i_k}\}$. بنابراین دنباله $\{y_i\}_{i=۰}^{۲k-۱}$ در یک رابطه

[۱] Reed-Solomon

بازگشتی[۲] خطی و همگن از مرتبه k صدق می‌کند:

$$\exists c_1, \dots, c_k \in \mathbb{R}: \quad y_i + \sum_{j=1}^{k} c_j y_{i-j} = 0 \qquad (k \le i \le 2k-1) \tag{۲۱-۵}$$

در واقع x_iهای ناصفر، قطب‌های تبدیل $\mathcal{Z}$ دنباله y_i خواهند بود (اگر دنباله را بطور نامتناهی می‌ساختیم). برای یافتن x_{i_j}ها، اولین گام، یافتن ضرایب دنباله بازگشتی y_i است (c_iها). برای این منظور از k-معادله و k-مجهول زیر استفاده می‌کنیم(مقدار $y_0 = \sum_{i=1}^{n} a_i$ از قبل معلوم است):

$$\begin{bmatrix} y_{k-1} & y_{k-2} & \cdots & y_0 \\ y_k & y_{k-1} & \cdots & y_1 \\ \vdots & \vdots & \ddots & \vdots \\ y_{2k-2} & y_{2k-3} & \cdots & y_{k-1} \end{bmatrix} \cdot \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = \begin{bmatrix} -y_k \\ -y_{k+1} \\ \vdots \\ -y_{2k-1} \end{bmatrix} \tag{۲۲-۵}$$

بنابراین با حل دستگاه فوق، ضرایب c_i بدست می‌آیند (حالتی که این دستگاه تکین باشد، در ادامه مورد بحث قرار خواهد گرفت). اکنون همانطور که گفته شد، مقادیر ناصفر بردار $\mathbf{x}_{n\times 1}$ (x_{i_j}ها) ریشه چندجمله‌ای با ضرایب c_i هستند:

$$q(z) = z^k + c_1 z^{k-1} + c_2 z^{k-2} + \cdots + c_k$$

$$\Rightarrow \quad q(x_{i_1}) = q(x_{i_2}) = \cdots = q(x_{i_k}) = 0 \tag{۲۳-۵}$$

بنابراین با بدست آوردن ریشه‌های چندجمله‌ای $q(z)$ (پس از بدست آمدن c_iها) مقادیر ناصفر بردار $\mathbf{x}_{n\times 1}$ حاصل می‌شوند (روش فوق شباهت زیادی به روش ELP در یافتن مکان‌های خطا در کدهای حقیقی [۷۱] و یافتن زمان‌های ناپیوستگی به کمک Annihilating Filter در مبحث نمونه‌برداری از سیگنال‌های با نرخ نوآوری محدود دارد [۹۳]). دقت کنید که در این روش، ابتدا مقادیر ناصفر بردار $\mathbf{x}_{n\times 1}$ محاسبه می‌شوند و سپس درپی مکان این مقادیر در بردار $\mathbf{x}_{n\times 1}$ خواهیم بود (در روش‌های رایج نمونه‌برداری فشرده، اول مکان‌ها تعیین می‌شوند و بعد مقادیر).

پس از بدست آوردن قطب‌های تبدیل $\mathcal{Z}$ دنباله y_i (x_{i_j}ها)، به راحتی می‌توان به کمک $y_0, \dots, y_{k-1}$ که همانند شرایط اولیه برای دنباله بازگشتی است، ضریب ثابت هر قطب را هم بدست آورد (d_j):

$$y_l = \sum_{j=1}^{k} d_j x_{i_j}^l \qquad , \qquad 0 \le l \le 2k-1 \tag{۲۴-۵}$$

---

[۲] Difference Equation

از طرفی این ضرایب (d_jها) که از حل دنباله بازگشتی حاصل شده‌اند باید همان a_iهای اولیه باشند که هنـگام نمونه‌برداری انتخاب شده‌اند. پس، از تطابق d_jها با a_iها می‌توانیم تشخیص دهیم هرکدام از ریشه‌های بدست آمده مربوط به کدام مکان در بردار $\mathbf{x}_{n\times 1}$ بوده‌اند. برای جلوگیری از ابهام در این تطابق، باید a_iها متمایز باشند (درصورت برابری a_j و a_i، معلوم نیست که x_i ناصفر بوده است یا x_j). بنابراین، در حالتی که دستگاه خطی مطرح شده تکین نباشد، جواب k-تنک به طور یکتا و قطعی بدست خواهد آمد.

اکنون حالتی مطرح می‌شود که دستگاه خطی تکین باشد:

$$\det \begin{bmatrix} y_{k-1} & y_{k-2} & \cdots & y_0 \\ y_k & y_{k-1} & \cdots & y_1 \\ \vdots & \vdots & \ddots & \vdots \\ y_{2k-2} & y_{2k-3} & \cdots & y_{k-1} \end{bmatrix} = 0 \qquad (۲۵-۵)$$

این امر بدین معناست که $\{y_i\}$ها در دنباله بازگشتی کوچکتری (نسبت به طول مفروض k) صدق می‌کنند. این اتفاق دو منشاء می‌تواند داشته باشد: اول این که تعداد مولفه‌های غیرصفر بردار $\mathbf{x}_{n\times 1}$ کمتر از k بوده و دیگری این که تعدادی از مولفه‌های غیرصفر با هم برابر بوده‌اند (در صورت برابری دو مقدار، ضرایب قطب‌ها با هم جمع می‌شوند و در نهایت به صورت یک قطب جلوه می‌کنند). مشابه قبل، ابتدا باید مقادیر ناصفر بدست آیند و سپس مکان‌ها و تکراری بودن و نبودن آن‌ها تعیین شود. در نتیجه، به یک دنباله بازگشتی مناسب نیاز است. برای این امر باید ابعاد دستگاه خطی را که از ابتدا k فرض شده بود، آنقدر کاهش دهیم تا به یک ماتریس غیرتکین دست یابیم:

$$\mathbf{A}_i = \begin{bmatrix} y_i & y_{i-1} & \cdots & y_0 \\ y_{i+1} & y_i & \cdots & y_1 \\ \vdots & \vdots & \ddots & \vdots \\ y_{2i} & y_{2i-2} & \cdots & y_i \end{bmatrix} \quad , \quad 0 \le i \le k-1 \qquad (۲۶-۵)$$

به بیان بهتر، باید بزرگترین iای را بیابیم که $\det \mathbf{A}_i \neq 0$. این i بیانگر تعداد مقادیر ناصفر متفاوت بین درایه‌های بردار $\mathbf{x}_{n\times 1}$ است. پس از بدست آمدن دنباله بازگشتی مرتبه i و متعاقباً مقادیر ناصفر بردار $\mathbf{x}_{n\times 1}$ به کمک $y_0,\ldots,y_i$ (شرایط اولیه دنباله بازگشتی)، ضرایب قطب‌ها (d_jها) حاصل می‌شوند. مشابهاً، هر d_j برابر با یک یا زیرجمعی از a_jهاست (هنگامی که چند درایه‌ی ناصفر از $\mathbf{x}_{n\times 1}$ برابر باشند، ضرایب a_j آن‌ها با هم جمع می‌شود تا d_j نهایی حاصل شود). اکنون برای تشخیص بدون ابهام مکان‌ها و تکرارها، باید a_jها به گونه‌ای باشند که هیچ دو زیرجمعی از آن‌ها برابر نشود. عکس این مطلب نیز صادق است؛ اگر هیچ دو زیرجمعی از a_jها با

یکدیگر برابر نباشند، بطور یکتا می‌توان از روی d_jها مکان‌ها و تکرارهای اولیه را تعیین کرد. یک انتخاب ساده برای a_jها، توان‌های ۲ است:

$$۱ \le j \le n: \quad a_j = ۲^j \tag{۲۷–۵}$$

پس با انتخاب مناسب a_jها، تمام بردارهای k–تنک اولیه به طور یقینی (در حالت بدون نویز) قابل بازسازی هستند.

اگر f تابعی یک به یک با $f(۰) = ۰$ باشد، در نمونه‌های غیرخطی تولید شده در (۵–۲۰) می‌توان به جای x_i از $f(x_i)$ استفاده کرد. در این صورت در روند بازسازی نیز به جای x_i، $f(x_i)$ به دست می‌آید و از آنجا که f یک به یک است، می‌توان به x_i دست پیدا کرد. شرط $f(۰) = ۰$ از این رو اهمیت دارد که تنک‌بودن بردار $[x_۱, \ldots, x_n]$ کماکان در $[f(x_۱), \ldots, f(x_n)]$ صادق است و بنابراین نتایج به دست آمده هنوز معتبر هستند. استفاده از تابع f در محدود کردن دامنه تغییرات نمونه‌ها بسیار موثر است. فرض کنید دامنه تغییرات هر x_i بین $[-۲, ۲]$ باشد و بردار $[x_۱, \ldots, x_n]$ شامل k درایه ناصفر باشد، در روش نمونه‌برداری غیرخطی توان‌های این مقادیر تا مرتبه $۲k-۱$ باید محاسبه شوند که محدوده تغییرات را به $[-۲^{۲k-۱}, ۲^{۲k-۱}]$ افزایش می‌دهد. اکنون اگر تابع $f(x) = e^{\frac{j\pi}{۲}x}$ مورد استفاده قرار گیرد، محدوده تغییرات $f^i(x)$ برای هر i ثابت است. از طرفی این تابع برای محدوده مورد نظر x_i یک به یک است. پس با استفاده از تابع f مناسب می‌توان مشکل محدوده تغییرات را حل کرد. به عنوان یک مثال غیرخطی با نمونه‌های حقیقی، نمونه‌های زیر را درنظر بگیرید (تغییرات x_i محدود به $[-M, M]$ است):

$$\begin{cases} y_۱ &= a_۱ \sin\left(\frac{\pi x_۱}{۲M}\right) + a_۲ \sin\left(\frac{\pi x_۲}{۲M}\right) + \cdots + a_n \sin\left(\frac{\pi x_n}{۲M}\right) \\ y_۲ &= a_۱ \sin\left(۲\frac{\pi x_۱}{۲M}\right) + a_۲ \sin\left(۲\frac{\pi x_۲}{۲M}\right) + \cdots + a_n \sin\left(۲\frac{\pi x_n}{۲M}\right) \\ &\vdots \\ y_{۲k-۱} &= a_۱ \sin\left((۲k-۱)\frac{\pi x_۱}{۲M}\right) + a_۲ \sin\left((۲k-۱)\frac{\pi x_۲}{۲M}\right) + \cdots + a_n \sin\left((۲k-۱)\frac{\pi x_n}{۲M}\right) \end{cases} \tag{۲۸–۵}$$

از آنجا که $\sin x = \frac{e^{jx} - e^{-jx}}{۲j}$، هر تابع سینوسی در نمونه‌های فوق را می‌توان به دو تابع نمایی مختلط تجزیه کرد، این موضوع نشان می‌دهد که دنباله $\{y_l\}_{l=۱}^{\infty}$ در یک رابطه بازگشتی مرتبه $۲k$ صدق می‌کند و در نتیجه اگر روشی مشابه قبل استفاده شود، به $۴k-۱$ نمونه احتیاج است. نشان می‌دهیم در این حالت با $۲k$ نمونه می‌توان بردار k–تنک را بازیابی کرد. با تعریف چندجمله‌ای $q(z)$ به صورت قبل داریم:

$$\forall ۱ \le l \le k: \quad q(z) = \sum_{i=۰}^{۲k} c_i z^i = \prod_{l=۱}^{k} \left(z - e^{j\frac{\pi x_{i_l}}{۲M}}\right)\left(z - e^{-j\frac{\pi x_{i_l}}{۲M}}\right) = \prod_{l=۱}^{k} \left(z^۲ - ۲z\cos\frac{\pi x_{i_l}}{۲M} + ۱\right) \tag{۲۹–۵}$$

در نتیجه $c_۰ = c_{۲k} = ۱$ و $q(z)$ یک چندجمله‌ای متقارن است:

$$c_i = c_{۲k-i} \tag{۵–۳۰}$$

پس

$$\forall l: \quad ۰ = \sum_{i=۰}^{۲k} c_i y_{l-i} = c_k y_{l-k} + \sum_{i=۰}^{k-۱} c_i (y_{l-i} + y_{l-۲k+i}) \tag{۵–۳۱}$$

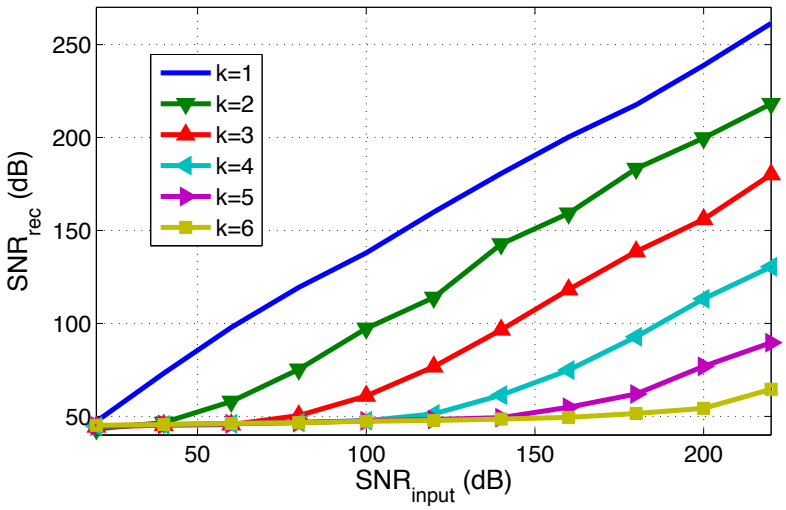
پس برای به دست آوردن c_iها می‌توانیم از حل دستگاه زیر بهره جوییم:

$$\begin{bmatrix} y_k + y_{۲-k} & y_{k-۱} + y_{۳-k} & \cdots & y_۲ + y_۰ & y_۱ \\ y_{k+۱} + y_{۳-k} & y_k + y_{۴-k} & \cdots & y_۳ + y_۱ & y_۲ \\ \vdots & \vdots & \ddots & \vdots & \\ y_{۲k-۱} + y_۱ & y_{۲k-۲} + y_۲ & \cdots & y_{k+۱} + y_{k-۱} & y_k \end{bmatrix} \begin{bmatrix} c_۱ \\ c_۲ \\ \vdots \\ c_k \end{bmatrix} = \begin{bmatrix} -(y_{k+۱} + y_{۱-k}) \\ -(y_{k+۲} + y_{۲-k}) \\ \vdots \\ -(y_{۲k} + y_۰) \end{bmatrix} \tag{۵–۳۲}$$

در اینجا نیز $y_۰ = ۰$ از پیش معلوم است و همچنین با توجه به تعریف نمونه‌ها داریم $y_{-l} = -y_l$؛ در نتیجه با داشتن $y_۱, \ldots, y_{k-۱}$، مقادیر $y_{-۱}, \ldots, y_{۱-k}$ معلوم خواهند بود. پس برای حل دستگاه (۵–۳۲) تنها $۲k$ نمونه از y_iها لازم است. با به دست آمدن ضرایب $c_۱, \ldots, c_k$، مشابه قبل چندجمله‌ای $q(z)$ را تشکیل می‌دهیم و پس از به دست آوردن ریشه‌های آن، مقادیر ناصفر بردار $[x_۱, \ldots, x_n]$ را آشکار می‌کنیم و با کمک ضرایب a_i و دنباله بازگشتی $\{y_l\}_l$، مکان این مقادیر نیز مشخص می‌شوند.

اهمیت k–تنک‌بودن بردار $\mathbf{x}$ در نمونه‌برداری غیرخطی بررسی شده تنها در وجود k مقدار غیرصفر مختلف در بین مقادیر بردار است؛ به عبارت بهتر، نکته اصلی تعداد مقادیر ناصفر نیست، بلکه تنوع در مقادیر ناصفر است. اگر تمامی مقادیر بردار $\mathbf{x}$ ناصفر باشند اما تنها k مقدار مختلف در بین درایه‌های آن یافت شود، دنباله $\{y_l\}_l$ همچنان در یک دنباله بازگشتی درجه k صدق خواهد کرد و مطابق مطالب بیان شده، این مقادیر و مکان‌های آن‌ها (در صورت مناسب انتخاب شدن ضرایب a_i) از روی نمونه‌ها قابل محاسبه‌اند. بنابراین، نمونه‌برداری غیرخطی نه تنها تعداد نمونه لازم و پیچیدگی محاسباتی عمل بازسازی را کاهش می‌دهد، بلکه بازسازی کامل را برای مجموعه بردارهای k–سطحی که تعمیمی از بردارهای k–تنک هستند تضمین می‌کند.

همان‌طور که در ابتدا ذکر شد، درخصوص نویز، به دلیل استفاده از عملگرهای غیرخطی، تحلیل عملکرد بسیار پیچیده است؛ البته در نویزهای کم، افت چندانی مشاهده نمی‌شود و در مورد عملکرد، می‌توان انتظار مشابهی نسبت به Annihilating Filter [۹۳] داشت.



شکل ۵-۱: کیفیت سیگنال بازسازی شده توسط نمونه‌های غیرخطی برای مرتبه‌های متفاوت تنک‌بودن.

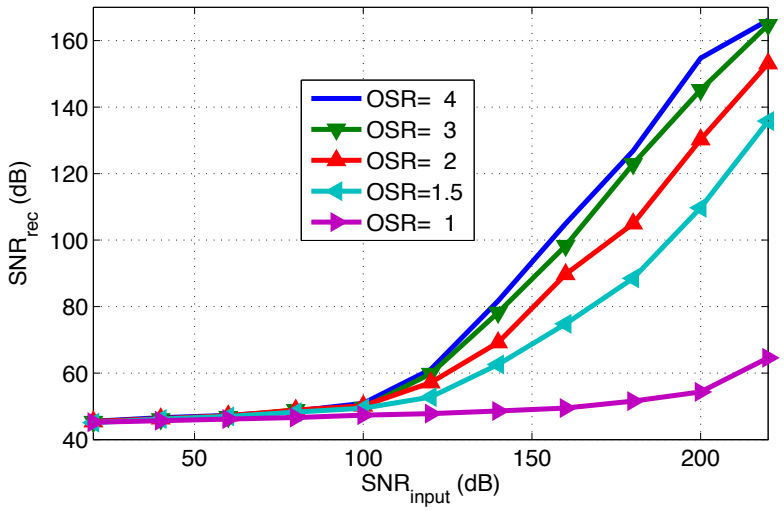
۵-۴ نتایج شبیه‌سازی

عملکرد نمونه‌برداری غیرخطی را با روش تابع سینوسی که مشکل محدوده تغییرات ندارد، بررسی می‌کنیم. برای این منظور بردارهای k-تنک با طول $n = ۵۰$ در نظر گرفته‌ایم که مکان‌های ناصفر به صورت تصادفی و هم‌احتمال از بین $\binom{۵۰}{k}$ حالت مختلف انتخاب می‌شوند. همچنین مقادیر غیرصفر با توزیع یکنواخت در $[-۲, ۲]$ اختیار می‌شوند. از آنجا که احتمال برابری دو مقدار ناصفر در بردار بسیار ناچیز است، نیازی به انتخاب $a_i = ۲^i$ نیست و برای این منظور از a_iهای تصادفی با توزیع گوسی ومیانگین صفر استفاده کرده‌ایم. برای بررسی اثر نویز، نمونه‌های غیرخطی تحت تاثیر نویز جمعی گوسی با واریانس‌های مشخص قرار گرفته‌اند تا SNRهای متفاوتی ایجاد شود.

شکل ۵-۱ کیفیت سیگنال بازسازی شده (SNR) را بر حسب SNR ورودی برای $k = ۱, \ldots, ۶$ نشان می‌دهد. برای به دست آوردن این نمودار، برای هر k، از $۲k$ نمونه استفاده شده است و نتایج برای ۵۰۰ شبیه‌سازی متفاوت میانگین‌گیری شده‌اند. شکل ۵-۱ نشان می‌دهد که با افزایش k، حساسیت نسبت به نویز افزایش می‌یابد و افت کیفیت مشهودتر می‌شود.

هنگامی که بیشتر از $۲k$ نمونه از y_iها در دسترس باشد، معادله (۵-۳۲) تبدیل به یک دستگاه فرا-معین[۳]

[۳] Over-determined



شکل ۵-۲: کیفیت سیگنال بازسازی شده توسط نمونه‌های غیرخطی به ازای $k = ۶$ در OSRهای متفاوت.

می‌شود و در نتیجه تاثیر نویز ورودی کاهش می‌یابد. در شکل ۵-۲ این اثر را برای $k = ۶$ بررسی کرده‌ایم. در این شکل منظور از OSR نسبت تعداد نمونه در دسترس به حداقل مورد نیاز $(۲k)$ است. مثلا، OSR= ۱٫۵ به معنای استفاده از $۳k$ نمونه است.

فصل ۶

دنباله‌های تصادفی فشرده‌پذیر

۶-۱ مقدمه

همان‌طور که در فصول قبل اشاره شد، در مبحث نمونه‌برداری فشرده، نمونه‌برداری و بازسازی سیگنال‌های تنک و یا در حالت کلی‌تر، فشرده‌پذیر، مورد بررسی قرار می‌گیرد. در فصل‌های گذشته، جزئیات مربوط به نحوه‌ی نمونه‌برداری و روش‌های بازسازی مورد مطالعه قرار گرفتند. در این فصل، به مبحث پایه‌ای‌تر فشرده‌پذیری یک دنباله می‌پردازیم. در واقع، تمامی استدلال‌هایی که تاکنون ارائه شده، بر این پایه استوارندکه بردار نمونه‌های یک سیگنال پیوسته، در یک پایه خاص، نمایش تنک و یا فشرده‌پذیر دارند. بردار نمونه‌ها همواره شامل متناهی تعداد است حال آن که فشرده‌پذیری و یا تنک‌بودن مربوط به خواص ذاتی سیگنال است و در نتیجه، توقع می‌رود که بتوان فشرده‌پذیری را حتی برای دنباله‌های نامتناهی نیز بررسی کرد. از آنجا که در کاربردهای عملی، سیگنال‌های واقعی را به کمک مدل‌های آماری بیان می‌کنند، در این فصل، مدل‌سازی آماری دنباله‌های نامتناهی فشرده‌پذیر مورد بررسی قرار می‌گیرد. مطالب این فصل برگرفته از [۸] می‌باشند.

در گذشته روند رایج در بررسی فشرده‌پذیری یک دنباله نامتناهی، استفاده از فضاهای Besov بود. این فضاها زیر فضاهایی از دنباله‌های با نرم ℓ_p محدود هستند که خطای تقریبی آن‌ها به کمک بهترین k جمله، به صورت k^{-c} با افزایش k افت می‌کند (c یک ثابت است) [۲۶، ۳۳]. در تعاریف جدید، فشرده‌پذیری به کمک نرم ℓ_p ضعیف بیان می‌شود: دنباله $\{x_i\}_{i=1}^{\infty}$ فشرده‌پذیر با نرم ℓ_p نامیده می‌شود اگر $R > 0$ موجود باشد که:

$$|x_i| \leq R i^{-\frac{1}{p}} \qquad (۱-۶)$$

با وجود نقاط قوت بسیار تعریف فوق، هنگامی که دنباله‌های تصادفی مورد نظر باشند، این تعریف کارآ نیست. برای روشن شدن این موضوع، فرض کنیم $\{x_i\}$ یک دنباله .i.i.d با توزیع احتمالی f_x است. در این صورت به غیر از چند حالت بدیهی، دنباله $\{x_i\}$ اصلاً میرا نخواهد بود و در نتیجه رابطه (۶-۱) صادق نیست. از طرف دیگر دنباله نمونه‌های تحقق‌های فرآیندهای تصادفی ایستان را می‌توان در حوزه نوآوری توسط دنباله‌های .i.i.d مدل کرد و در نتیجه فشرده‌پذیری سیگنال، منوط به فشرده‌پذیری دنباله نوآوری آن است. از آن‌جا که تعبیر آماری روش موفق Basis Pursuit (کمینه کردن نرم ℓ_1) به صورت تخمین گر MAP توزیع احتمال لاپلاس را برای ورودی .i.i.d پیشنهاد می‌کند، در بعضی موارد توزیع احتمال لاپلاس را به طور اشتباه یک توزیع فشرده‌پذیر قلمداد می‌کنند [۹۵، ۷۷]. این اشتباه در چندین مقاله مورد نقد قرار گرفته است [۲۷، ۲۸، ۹۹]. بررسی توزیع‌های فشرده‌پذیر، اولین بار در [۲۷] به صورت نسبتاً دقیق انجام گرفته‌است. در این مقاله به کمک تقریب‌های توزیع احتمال دنباله ترتیب‌ها[1] تعدادی از توزیع احتمال‌ها مثل Generalized Pareto و Student's t به عنوان توزیع فشرده‌پذیر معرفی شده‌اند. نقطه ضعف اصلی این مقاله، استفاده از تقریب در تعریف هاست که تعاریف را نادقیق می‌کند. علاوه بر این، نتایج به گونه‌ای است که بررسی فشرده‌پذیری توزیع‌های احتمال باید به صورت مورد به مورد انجام پذیرد؛ به عبارت بهتر، نمی‌توان فشرده‌پذیری را به صورت کامل برای یک دسته از توزیع‌ها نشان داد. در این فصل، به کمک نتایج حدی در مورد توزیع احتمال دنباله‌های مرتب، فشرده‌پذیری را به طور دقیق تعریف می‌کنیم و به کمک چند قضیه، دسته بزرگی از توزیع احتمال‌های فشرده‌پذیر را معرفی می‌کنیم.

۶-۲ دنباله‌های نامتناهی فشرده‌پذیر

برای بررسی فشرده‌پذیری دنباله‌های تصادفی، ابتدا از دنباله‌های یقینی شروع می‌کنیم. همانطور که قبلاً اشاره شد، می‌خواهیم مفهوم فشرده‌پذیری را بدون استفاده از میرایی دنباله تعریف کنیم تا بتوان این تعریف را به دنباله‌های تصادفی نیز تعمیم داد. به طور شهودی، یک دنباله متناهی فشرده‌پذیر است هرگاه درصد بالایی از انرژی کل دنباله در تعداد کمی از جملات متمرکز شده باشد. اکنون یک دنباله نامتناهی را فشرده‌پذیر می‌نامیم هرگاه تمام زیر دنباله‌های منقطع[2] آن فشرده‌پذیر باشد. برای تعریف دقیق این مفهوم، ابتدا نمادهای زیر را تعریف می‌کنیم:

تعریف ۶-۱ فرض کنید $\{a_i\}_{i=1}^n$ دنباله‌ای از اعداد حقیقی (مختلط) باشد و از لحاظ قدر مطلق، دنباله مرتب

[1] Order Statistics

[2] Truncated

شده آن $\{a_{ni}\}_{i=۱}^{n}$ باشد ($|a_{n۱}| \geq \cdots \geq |a_{nn}|$). برای $p > ۰$ و $k \leq n$ تعریف می‌کنیم:

$$\sigma_p\big(k;\{a_i\}_{i=۱}^{n}\big) = \Big(|a_{n۱}|^p + \cdots + |a_{nk}|^p\Big)^{\frac{۱}{p}} \qquad (۶\text{-}۲)$$

به عبارت بهتر، $\sigma_p(k;\{a_i\}_{i=۱}^{n})$ برابر است با نرم ℓ_p بهترین تقریب k جمله‌ای از دنباله.

تعریف ۶-۲ برای دنباله $\{a_i\}_{i=۱}^{n}$ که جملات آن همگی صفر نیستند و $۰ \leq r \leq ۱$ و $p > ۰$ مفروض، تعریف می‌کنیم:

$$\kappa_p\big(r;\{a_i\}_{i=۱}^{n}\big) = \min\left\{k \;\middle|\; \frac{\sigma_p\big(k;\{a_i\}_{i=۱}^{n}\big)}{\sigma_p\big(n;\{a_i\}_{i=۱}^{n}\big)} \geq r\right\} \qquad (۶\text{-}۳)$$

به بیان ساده‌تر $\kappa_p(.)$ نشان‌دهنده حداقل تعداد جمله مورد نیاز برای بیان کسر r از کل انرژی دنباله است.

حال فشرده‌پذیری یک دنباله را چنین تعریف می‌کنیم:

تعریف ۶-۳ دنباله $\{a_i\}_{i\in\mathbb{N}}$ از اعداد حقیقی (مختلط) را ℓ_p-فشرده‌پذیر می‌نامیم هر گاه :

$$\forall\, ۰ \leq r < ۱ : \quad \lim_{n\to\infty} \frac{\kappa_p\big(r;\{a_i\}_{i=۱}^{n}\big)}{n} = ۰ \qquad (۶\text{-}۴)$$

تعریف فوق بیان می‌دارد که در یک دنباله فشرده‌پذیر، برای دستیابی به هر کسری از کل انرژی، تنها تعداد اندکی از کل جملات مورد نیاز هستند. برای بررسی این تعریف، مثال $i \in \mathbb{N},\ \ a_i = \frac{۱}{\alpha^i}$ را در نظر می‌گیریم که $|\alpha| \neq ۱$. برای $۰ \leq r < ۱$ دلخواه قرار می‌دهیم $k = \lceil -\frac{\ln(۱-r^p)}{p|\ln|\alpha||} \rceil$. در این صورت داریم:

$$\begin{aligned}\frac{\sigma_p\big(k;\{a_i\}_{i=۱}^{n}\big)}{\sigma_p\big(n;\{a_i\}_{i=۱}^{n}\big)} &= \left(\frac{\sum_{i=۱}^{k} e^{-ip\left|\ln|\alpha|\right|}}{\sum_{i=۱}^{n} e^{-ip\left|\ln|\alpha|\right|}}\right)^{\frac{۱}{p}} \geq \left(\frac{\sum_{i=۱}^{k} e^{-ip\left|\ln|\alpha|\right|}}{\sum_{i=۱}^{\infty} e^{-ip\left|\ln|\alpha|\right|}}\right)^{\frac{۱}{p}} \\ &= \Big(۱ - e^{-\left|\ln|\alpha|\right|kp}\Big)^{\frac{۱}{p}} \geq r \end{aligned} \qquad (۶\text{-}۵)$$

در نتیجه $\kappa_p\big(r;\{a_i\}_{i=۱}^{n}\big) \leq k$. از آنجا که k فقط به r بستگی دارد و نه n، واضح است که:

$$\lim_{n\to\infty} \frac{\kappa_p\big(r;\{a_i\}_{i=۱}^{n}\big)}{n} = ۰ \qquad (۶\text{-}۶)$$

و در نتیجه دنباله $\{a_i\}_{i\in\mathbb{N}}$ برای هر $p > ۰$، ℓ_p-فشرده‌پذیر است. مشابه استدلال فوق، می‌توان نشان داد که دنباله‌های با نرم ℓ_p محدود، ℓ_p-فشرده‌پذیر هم هستند. نکته جالب توجه در مثال فوق این است که در حالت $|\alpha| > ۱$، دنباله $\{a_i\}$ نه تنها میرا نیست، بلکه به صورت نمایی به سمت بی‌نهایت میل می‌کند.

حال به بررسی خواص دنباله‌های فشرده‌پذیر می‌پردازیم. یکی از مهم‌ترین خواص این دنباله‌ها، تودرتو بودن آن‌هاست؛ یعنی اگر دنباله‌ای ℓ_p–فشرده‌پذیر باشد، برای هر $q \geq p$ و ℓ_q–فشرده‌پذیر نیز هست. برای اثبات این موضوع از لم زیر کمک می‌گیریم:

لم ۶–۱ برای دنباله دلخواه $\{a_i\}$ و اعداد صحیح $k < n$، $\left(\frac{\sigma_p(k;\{a_i\}_{i=1}^n)}{\sigma_p(n;\{a_i\}_{i=1}^n)}\right)^p$ نسبت به p تابعی صعودی است.

اثبات: برای اندیس‌های $1 \leq i \leq j \leq n$ و اعداد $0 < p \leq q$ داریم:

$$|a_{ni}| \geq |a_{nj}| \quad \Rightarrow \quad |a_{ni}|^{q-p} \geq |a_{nj}|^{q-p}$$

$$\Rightarrow \quad |a_{ni}|^q |a_{nj}|^p \geq |a_{ni}|^p |a_{nj}|^q \tag{۶–۷}$$

اکنون با جمع کردن تمام این نامساوی‌ها برای $1 \leq i \leq k$ و $k+1 \leq j \leq n$ داریم:

$$\begin{aligned} \sum_{i=1}^{k} \sum_{j=k+1}^{n} |a_{ni}|^q |a_{nj}|^p &\geq \sum_{i=1}^{k} \sum_{j=k+1}^{n} |a_{ni}|^p |a_{nj}|^q \\ \Rightarrow \frac{\sum_{i=1}^{k} |a_{ni}|^q}{\sum_{i=k+1}^{n} |a_{ni}|^q} &\geq \frac{\sum_{i=1}^{k} |a_{ni}|^p}{\sum_{i=k+1}^{n} |a_{ni}|^p} \\ \Rightarrow \frac{\sum_{i=1}^{k} |a_{ni}|^q}{\sum_{i=1}^{n} |a_{ni}|^q} &\geq \frac{\sum_{i=1}^{k} |a_{ni}|^p}{\sum_{i=1}^{n} |a_{ni}|^p} \end{aligned} \tag{۶–۸}$$

که اثبات را به پایان می‌رساند. ■

لم ۶–۱ نتیجه می‌دهد که $\kappa_p(r^{\frac{1}{p}}, \{a_i\}_{i=1}^n)$ تابعی نزولی نسبت به p است. پس اگر حد مورد نظر در تعریف ۶–۳ برای یک مقدار p به صفر میل کند، این حد برای تمامی مقادیر $q \geq p$ نیز به صفر میل خواهد کرد. پس هر دنباله نامتناهی یا فشرده ناپذیر است و یا برای p از جایی به بعد ℓ_p–فشرده‌پذیر است. کوچکترین مقدار p که برای آن دنباله ℓ_p–فشرده‌پذیر است، معیاری از فشرده‌پذیری ارائه می‌دهد. یعنی دنباله ℓ_p– فشرده‌پذیر با p کمتر، فشرده‌پذیرتر تلقی می‌شود.

۶–۳ فشرده‌پذیری دنباله‌های تصادفی

برای تعمیم مفهوم فشرده‌پذیری به دنباله‌های تصادفی، لازم است که احتمال پیشامدهای مختلف نیز در تعریف گنجانده شود. به طور شهودی، یک دنباله تصادفی فشرده‌پذیر است اگر تحقق‌های آن با احتمال به دلخواه زیاد (نزدیک به یک) فشرده‌پذیر باشند. برای بیان دقیق این تعریف، ابتدا تعریف ۶–۲ را مجدداً به کمک توزیع احتمال بازنویسی می‌کنیم:

تعریف ۶-۴ فرض کنید $\{x_i\}_{i\in\mathbb{N}}$ یک دنباله *i.i.d.* از متغیرهای تصادفی با توزیع احتمال f_x باشد. همچنین $0 \leq r, \delta \leq 1$ و $p \geq 0$ اعدادی دلخواه هستند. تعریف می‌کنیم:

$$\tilde{\kappa}_p(r, \delta, n; f_x) = \min\left\{k \;\middle|\; P\Big(\frac{\sigma_p\big(k; \{x_i\}_{i=1}^n\big)}{\sigma_p\big(n; \{x_i\}_{i=1}^n\big)} \geq r\Big) \geq \delta\right\} \tag{۶-۹}$$

در حقیقت $\tilde{\kappa}_p(.)$ همان نقش $\kappa_p(.)$ در دنباله‌های یقینی را ایفا می‌کند با این تفاوت که به جای نامساوی قطعی، نامساوی با احتمال δ برقرار است. به این ترتیب در تعریف $\tilde{\kappa}_p$ دو کران پایین دخالت دارند. یکی بیان‌گر حداقل درصد انرژی لازم (r) و دیگری حداقل احتمال وقوع پیشامد (δ) است. از آن‌جا که تمام خواص یک دنباله i.i.d. براساس توزیع احتمال آن تعیین می‌شود، فشرده‌پذیری یک دنباله تصادفی را به جای دنباله، به خود توزیع احتمال نسبت می‌دهیم. تعریف زیر مشابه حالت غیرتصادفی، فشرده‌پذیری تصادفی (برای توزیع احتمال) را براساس رفتار حدی $\tilde{\kappa}_p$ بیان می‌کند:

تعریف ۶-۵ توزیع احتمال f_x را ℓ_p-فشرده‌پذیر می‌نامیم هر گاه:

$$\forall\, 0 \leq r, \delta < 1, \quad \lim_{n\to\infty} \frac{\tilde{\kappa}_p(r, \delta, n; f_x)}{n} = 0 \tag{۶-۱۰}$$

دقت کنید که حالات $\delta = 1, r = 1$ در رابطه فوق درنظر گرفته نشده‌اند. در بردارهای با طول متناهی، براساس این که بهترین تقریب k-جمله از لحاظ انرژی نمایش دقیق و یا تقریبی از بردار ایجاد می‌کند، آن را تنک یا فشرده‌پذیر می‌نامند. در این‌جا نیز می‌توان دنباله تصادفی تنک را به طور مشابهی تعریف کرد:

تعریف ۶-۶ توزیع احتمال f_x را ℓ_p-تنک می‌نامیم هر گاه ℓ_p-فشرده‌پذیر باشد و در هر تحقق از دنباله‌های *i.i.d.* با این توزیع احتمال، کل انرژی دنباله در کسری از جملات متمرکز باشد:

$$\forall\, p,\ 0 \leq \delta < 1, \quad \lim_{n\to\infty} \frac{\tilde{\kappa}_p(1, \delta, n; f_x)}{n} < 1 \tag{۶-۱۱}$$

به زبان ساده‌تر، تنها فرق فشرده‌پذیری و تنک‌بودن، رفتار دنباله در حالت $r = 1$ است. در یک دنباله تنک توقع داریم همواره کسری از جملات صفر باشند و در انرژی کل مشارکت نکنند. در واقع (۶-۱۱) بیان می‌کند که $P(x = 0) = \pi_0 > 0$ و یا به عبارت دیگر توزیع احتمال f_x در نقطه $x = 0$ شامل یک تابع ضربه است. از سوی دیگر اگر $P(x = 0) = \pi_0 > 0$، با احتمال بسیار بالا کل انرژی یک دنباله nتایی در $n(1 - \pi_0)$ جمله قرار می‌گیرد. با وجود شرط اضافه در تعریف تنک‌بودن، این توزیع‌ها به راحتی براساس توزیع‌های فشرده‌پذیر

مشخص می‌شوند. فرض کنید f_x یک توزیع فشرده‌پذیر باشد، برای تولید یک دنباله تنک می‌توان ابتدا به کمک f_x یک دنباله فشرده‌پذیر ساخت و سپس با احتمال π_0 هریک از این جملات را صفر کرد و یا ثابت نگه داشت. به راحتی می‌توان نشان داد که دنباله‌ی حاصل، تنک است. شایان ذکر است که توزیع‌های فشرده‌پذیر و یا تنک، ساختاری تودرتو مشابه آن چه در لم ۶-۱ برای ساختارهای غیرتصادفی نشان داده شد، دارند: یک توزیع ℓ_p-فشرده‌پذیر (ℓ_p-تنک) برای هر $q \geq p$ ℓ_q-فشرده‌پذیر (ℓ_q-تنک) نیز هست. همچنین لم زیر رابطه بین توزیع‌های ℓ_p-فشرده‌پذیر و ℓ_1-فشرده‌پذیر را نشان می‌دهد:

لم ۶-۲ توزیع احتمال متغیر تصادفی x از نوع ℓ_p-فشرده‌پذیر است اگر و تنها اگر، توزیع احتمال متغیر $y = |x|^p$ ℓ_1-فشرده‌پذیر باشد.

اثبات: با توجه به این نکته که نرم ℓ_p دنباله $\{x_i\}$ با نرم ℓ_1 دنباله $\{y_i\}$ معادل است، لم ثابت می‌شود.

۶-۴ تشخیص توزیع احتمال‌های فشرده‌پذیر

در ادامه توسط دو قضیه، دسته بزرگی از توزیع‌های فشرده‌پذیر را مشخص می‌کنیم. ابتدا در قضیه اول نشان می‌دهیم که توزیع‌هایی که به صورت نمایی میرا می‌شوند فشرده ناپذیرند. در قضیه دوم نیز نشان می‌دهیم که اگر میرایی به صورت چندجمله‌ای باشد، توزیع فشرده‌پذیر است. با توجه به لم ۶-۲ برای شناسایی توزیع‌های فشرده‌پذیر، فقط لازم است توزیع‌های ℓ_1-فشرده‌پذیر را شناسایی کنیم. از این رو در ادامه تنها بر روی دنباله‌های ℓ_1-فشرده‌پذیر متمرکز می‌شویم.

قضیه ۶-۱ اگر توزیع احتمال f_x به گونه‌ای باشد که برای یک $\gamma > 0$، $\mathscr{E}\{e^{\gamma|x|}\}$ محدود باشد، این توزیع ℓ_1-فشرده‌پذیر نیست.

اثبات: فرض کنید $\{x_i\}$ دنباله‌ای i.i.d. با توزیع احتمال f_x باشد و قرار دهید $y_i = |x_i|$. به وضوح دنباله $\{y_i\}$ نیز i.i.d. خواهد بود اما با توزیع $f_y(y) = \big(f_x(y) + f_x(-y)\big)u(y)$. همچنین فرض کنید μ_y و σ_y^2 به ترتیب میانگین و واریانس توزیع f_y باشند (از آن‌جا که $\mathscr{E}_x\{e^{\gamma|x|}\}$ محدود است، μ_y و σ_y نیز محدودند). برای $0 < r < 1$ دلخواه، فرض کنید m عدد مثبت به اندازه کافی بزرگ است که $\mathscr{E}_y\{e^{\gamma y}\} < e^{m\gamma r\mu_y}(3m)^{-1}$ (γ همان مقدار مثبتی است که برای آن $\mathscr{E}\{e^{\gamma y}\}$ محدود است). برای بررسی فشرده‌پذیری پیشامدهای زیر را تعریف

می‌کنیم:

$$\begin{cases} EV_1: & \sigma_1(k;\{y_i\}_{i=1}^{mk}) \geq r\sum_{i=1}^{mk} y_i \\ EV_2: & \sum_{i=1}^{mk} y_i \geq mk\mu_y - (mk)^{0.75}\sigma_y \\ EV_3: & \sum_{i=1}^{mk} y_i < mk\mu_y - (mk)^{0.75}\sigma_y \end{cases} \quad (۱۲-۶)$$

در نتیجه:

$$\begin{aligned} P(EV_1) &= P(EV_1|EV_2)P(EV_2) + P(EV_1|EV_3)P(EV_3) \leq P(EV_1|EV_2) + P(EV_3) \\ &\leq P\Big(\sigma_1(k;\{y_i\}_{i=1}^{mk}) \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)\Big) + P\Big(\sum_{i=1}^{mk} y_i < mk\mu_y - (mk)^{0.75}\sigma_y\Big) \end{aligned} \quad (۱۳-۶)$$

از آن‌جا که y_iها مستقل‌اند (i.i.d.) داریم:

$$\begin{cases} \mathscr{E}_y\{\sum_{i=1}^{mk} y_i\} = mk\mu_y \\ \mathrm{Var}\{\sum_{i=1}^{mk} y_i\} = mk\sigma_y^2 \end{cases} \quad (۱۴-۶)$$

بنابراین با کمک نامساوی چبیشف[۳] بدست می‌آید:

$$P\Big(\sum_{i=1}^{mk} y_i < mk\mu_y - (mk)^{0.75}\sigma_y\Big) \leq \frac{1}{\sqrt{mk}} \quad (۱۵-۶)$$

علاوه بر این، $\sigma_1(k;\{y_i\}_{i=1}^{mk})$ توسط یکی از $\binom{mk}{k}$ حالت انتخاب k جمله از کل mk جمله بدست آمده است.

پس:

$$P\Big(\sigma_1(k;\{y_i\}_{i=1}^{mk}) \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)\Big) \leq \binom{mk}{k} P\Big(\sum_{i=1}^{k} y_i \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)\Big) \quad (۱۶-۶)$$

از آن‌جا که $\mathscr{E}\{\sum_{i=1}^{k} y_i\} = k\mu_y$، احتمال برقراری $\sum_{i=1}^{k} y_i \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)$ جزء احتمالات انتهایی[۴] طبقه‌بندی می‌شود. اکنون نشان می‌دهیم این احتمال نسبت به k به صورت نهایی میرا می‌شود:

$$P\Big(\sum_{i=1}^{k} y_i \geq T\Big) = P\big(e^{\gamma\sum_{i=1}^{k} y_i} \geq e^{\gamma T}\big) \leq \frac{\mathscr{E}\{e^{\gamma\sum_{i=1}^{k} y_i}\}}{e^{\gamma T}} = \big(\mathscr{E}\{e^{\gamma y}\}e^{-\gamma\frac{T}{k}}\big)^k < \Big(\frac{e^{\gamma(rm\mu_y - \frac{T}{k})}}{3m}\Big)^k \quad (۱۷-۶)$$

که اولین نامساوی در رابطه فوق توسط نامساوی مارکف[۵] بدست آمده است. حال به استناد $\binom{mk}{k} < (em)^k$ و نتایج بدست آمده در (۶-۱۶) و (۶-۱۷) داریم:

$$P\Big(\sigma_1(k;\{y_i\}_{i=1}^{mk}) \geq r\big(mk\mu_y - (mk)^{0.75}\sigma_y\big)\Big) < \Big(\frac{e}{3}e^{\gamma r\sigma_y m^{0.75}k^{-0.25}}\Big)^k \quad (۱۸-۶)$$

[۳] Chebychev
[۴] Tail Probability
[۵] Markov

که با در نظر گرفتن (۶-۱۳) و نتیجه می‌دهد:

$$P(EV_1) = P\left(\frac{\sigma_1(k;\{x_i\}_{i=1}^{mk})}{\sigma_1(mk;\{x_i\}_{i=1}^{mk})} \geq r\right) < \frac{1}{\sqrt{mk}} + \left(\frac{e}{3}e^{\gamma r \sigma_y m^{0.75} k^{-0.25}}\right)^k \qquad (۶-۱۹)$$

به راحتی می‌توان بررسی کرد که برای یک m ثابت، هنگامی که $k \to \infty$ کران بالایی فوق و در نتیجه احتمال $P(EV_1)$ به صفر میل می‌کنند. از این رو برای این که احتمال $P(EV_1)$ از یک کران پایین دلخواه مانند δ بزرگتر باشد، به بیش از k جمله از کل mk جمله دنباله نیازمندیم. پس:

$$\lim_{k\to\infty} \frac{\tilde{\kappa}_p(r,\delta,km;f_x)}{km} \geq \frac{1}{m} \qquad (۶-۲۰)$$

که m تنها به r بستگی دارد و نه k. در نتیجه، توزیع احتمال f_x، ℓ_1-فشرده‌پذیر نیست. ■

قضیه ۶-۱ نشان می‌دهد که توزیع‌هایی نظیر لاپلاس، گاما و گوسی که میرایی نمایی دارند، ℓ_1-فشرده‌پذیر نیستند. در ادامه نشان می‌دهیم که در صورت میرایی از مرتبه چندجمله‌ای، توزیع احتمال فشرده‌پذیر است. پیش از بیان این نتیجه، ابتدا به بیان پیش نیازهایی در نظریه احتمال می‌پردازیم:

براساس قضیه معروف حد مرکزی[۶] اگر $\{x_i\}_{i\in\mathbb{N}}$ دنباله‌ای i.i.d. از متغیرهای تصادفی با میانگین (μ_x) و واریانس (σ_x^2) محدود باشد، زیرجمع‌های نرمالیزه شده آن‌ها به فرم $n^{-0.5}\sum_{i=1}^{n}(x_i-\mu_x)$ در حد، به سمت متغیرتصادفی گوسی با میانگین صفر و واریانس σ_x^2 میل می‌کند. علاوه بر این، جمع دو متغیرتصادفی مستقل گوسی مجدداً متغیر گوسی است؛ به بیان دیگر، توزیع گوسی نسبت به عمل جمع پایدار[۷] است. نکته جالب آن است که مجموعه توزیع‌های پایدار، تنها محدود به توزیع‌های گوسی نمی شود. دسته‌ای از توزیع‌ها معروف به α-پایدار موجودند که توسط پارامتر $0 < \alpha \leq 2$ اندیس‌گذاری می‌شوند و همگی نسبت به عمل جمع پایدارند؛ حالت خاص $\alpha = 2$ مربوط به توزیع گوسی است. به غیر از حالت $\alpha = 2$ (توزیع گوسی)، میرایی این توزیع‌ها به صورت $|t|^{-(\alpha+1)}$ است و در نتیجه ممان‌های آن‌ها تنها تا مرتبه α متناهی است. مثلاً هیچ کدام از آن‌ها واریانس متناهی ندارند [۸۲]. مشابه حالت گوسی، توزیع حدی زیرجمع‌های نرمالیزه شده (با $n^{-\frac{1}{\alpha}}$) دسته خاصی از توزیع احتمال‌ها به صورت α-پایدار خواهد بود؛ به مجموعه این توزیع احتمال‌ها حوزه جذب[۸] توزیع α-پایدار می‌گویند.

---

[۶] Central Limit Theorem

[۷] Stable

[۸] Domain of Attraction

برای شناسایی حوزه جذب در حالت $۰ < \alpha < ۲$، فرض کنید x یک متغیرتصادفی باشد و تعریف کنید $y = |x|$ و $G(t) = P(y > t)$. طبق نتیجه‌ای در نظریه احتمال [۴۵] می‌دانیم که توزیع احتمال x در حوزه جذب یک توزیع α-پایدار است اگر و فقط اگر:

۱. تابع $h(t) = t^{\alpha}G(t)$ تغییرات کندی در $t \to \infty$ داشته باشد، به بیان دیگر، برای هر $c > ۰$ حد $\frac{h(ct)}{h(t)}$ وقتی t به سمت بی‌نهایت میل می‌کند، برابر یک باشد؛

۲. حد $\lim_{t\to\infty} \frac{P(x>t)}{G(t)}$ موجود باشد.

همانطور که اشاره شد، به دلیل میرایی چندجمله‌ای در توزیع‌های متعلق به حوزه جذب توزیع‌های α-پایدار، این توزیع‌ها از دیدگاه فشرده‌پذیری گزینه‌های مناسبی به شمار می‌روند. برای اثبات این موضوع ابتدا قضیه زیر را که در [۶۵] اثبات شده است، بیان می‌کنیم:

قضیه ۶-۲ فرض کنید $\{\eta_i\}$ دنباله‌ای *i.i.d.* از متغیرهای تصادفی با توزیع احتمال نمایی استاندارد (میانگین واحد) باشند و فرض کنید $\Gamma_i = \sum_{j=۱}^{i} \eta_j$. اگر توزیع احتمال دنباله $\{x_i\}$ متعلق به حوزه جذب توزیع α-پایدار باشد داریم:

$$\lim_{n\to\infty} a_n^{-۱}\big(y_{n۱}, \ldots, y_{nn}, ۰, ۰, \ldots\big) =_d \big(\Gamma_۱^{-\frac{۱}{\alpha}}, \Gamma_۲^{-\frac{۱}{\alpha}}, \ldots\big) \qquad (۲۱\text{-}۶)$$

که در آن $=_d$ به معنای همگرایی در توزیع است و $\{y_{ni}\}_{i=۱}^{n}$ دنباله مرتب شده (نزولی) $\{|x_i|\}_{i=۱}^{n}$ را نشان می‌دهید. همچنین:

$$a_n = \lim_{t\to\infty} \frac{G^{-۱}\big(\frac{۱}{nt^{\alpha}}\big)}{t} \qquad (۲۲\text{-}۶)$$

نکته در خور توجه قضیه ۶-۲ این است که توزیع‌های حدی دنباله مرتب شده دسته‌ای از توزیع‌ها را می‌توان تنها با بررسی توزیع نمایی استاندارد بدست آورد. اکنون ابزار لازم برای نشان دادن فشرده‌پذیری دسته بزرگی از توزیع‌ها را در اختیار داریم:

قضیه ۶-۳ اگر توزیع احتمال متغیر x که با f_x نمایش داده می‌شود متعلق به حوزه جذب یک توزیع α-پایدار با $۰ < \alpha < ۱$ باشد، در این صورت f_x $\ell_۱$-فشرده‌پذیر است.

اثبات: برای اثبات فشرده‌پذیری از احتمال بیان شده در تعریف ۶-۴ شروع می‌کنیم. از آن‌جا که $\sigma_۱\big(k;\{x_i\}_{i=۱}^n\big)=\sigma_۱\big(k;\{y_i\}_{i=۱}^n\big)$ ($y_i=|x_i|$ که) داریم:

$$P\Big(\frac{\sigma_۱\big(k;\{x_i\}_{i=۱}^n\big)}{\sigma_۱\big(n;\{x_i\}_{i=۱}^n\big)}\geq r\Big)=P\Big(\frac{a_n^{-۱}\big(y_{n(k+۱)}+\cdots+y_{nn}\big)}{a_n^{-۱}\big(y_{n۱}+\cdots+y_{nn}\big)}\leq ۱-r\Big) \qquad (۲۳-۶)$$

در تعریف فشرده‌پذیری ما به دنبال رفتار حدی احتمال فوق هستیم. طبق قضیه ۶-۲ در حالت حدی می‌توانیم $a_n^{-۱}y_{ni}$ را با $\Gamma_i^{-\frac{۱}{\alpha}}$ جایگذاری کنیم (این موضوع را به طور دقیق‌تر در انتهای اثبات نشان می‌دهیم). در نتیجه داریم:

$$P\Big(\frac{\sigma_۱\big(k;\{x_i\}_{i=۱}^n\big)}{\sigma_۱\big(n;\{x_i\}_{i=۱}^n\big)}\geq r\Big)\sim P\Big(\frac{\Gamma_{k+۱}^{-\frac{۱}{\alpha}}+\cdots+\Gamma_n^{-\frac{۱}{\alpha}}}{\Gamma_۱^{-\frac{۱}{\alpha}}+\cdots+\Gamma_n^{-\frac{۱}{\alpha}}}\leq ۱-r\Big) \qquad (۲۴-۶)$$

دقت کنید :

$$\begin{aligned}\frac{\Gamma_{k+۱}^{-\frac{۱}{\alpha}}+\cdots+\Gamma_n^{-\frac{۱}{\alpha}}}{\Gamma_۱^{-\frac{۱}{\alpha}}+\cdots+\Gamma_n^{-\frac{۱}{\alpha}}} &= \frac{k^{۲-\alpha}\big(\frac{\Gamma_۱}{\Gamma_{k+۱}}\big)^{\frac{۱}{\alpha}}}{۱+\big(\frac{\Gamma_۱}{\Gamma_۲}\big)^{\frac{۱}{\alpha}}+\cdots+\big(\frac{\Gamma_۱}{\Gamma_n}\big)^{\frac{۱}{\alpha}}}\times\frac{۱+\big(\frac{\Gamma_{k+۱}}{\Gamma_{k+۲}}\big)^{\frac{۱}{\alpha}}+\cdots+\big(\frac{\Gamma_{k+۱}}{\Gamma_n}\big)^{\frac{۱}{\alpha}}}{k^{۲-\alpha}}\\ &\leq \underbrace{k^{۲-\alpha}\big(\frac{\Gamma_۱}{\Gamma_{k+۱}}\big)^{\frac{۱}{\alpha}}}_{A_k}\underbrace{\frac{۱+\big(\frac{\Gamma_{k+۱}}{\Gamma_{k+۲}}\big)^{\frac{۱}{\alpha}}+\cdots+\big(\frac{\Gamma_{k+۱}}{\Gamma_n}\big)^{\frac{۱}{\alpha}}}{k^{۲-\alpha}}}_{B_{k,n}} \qquad (۲۵-۶)\end{aligned}$$

که نتیجه می‌دهد:

$$\begin{aligned}P\Big(\frac{\Gamma_{k+۱}^{-\frac{۱}{\alpha}}+\cdots+\Gamma_n^{-\frac{۱}{\alpha}}}{\Gamma_۱^{-\frac{۱}{\alpha}}+\cdots+\Gamma_n^{-\frac{۱}{\alpha}}}\leq ۱-r\Big) &\geq ۱-P\big(A_kB_{k,n}\geq ۱-r\big)\geq ۱-P\big(A_k>۱\big)-P\big(B_{k,n}>۱-r\big)\\ &\geq ۱-\mathscr{E}\{A_k\}-\frac{۱}{۱-r}\mathscr{E}\{B_{k,n}\} \qquad (۲۶-۶)\end{aligned}$$

آخرین نامساوی با استفاده از نامساوی مارکف حاصل شده است. در ادامه اثبات نشان می‌دهیم هر دوی $\mathscr{E}\{A_k\}$ و $\mathscr{E}\{B_{k,n}\}$ در حالت حدی $k\to\infty$ و $n>k+۲$ به صفر میل می‌کنند.

برای بدست آوردن حد $\mathscr{E}\{A_k\}$ وقتی $k\to\infty$، ابتدا لم زیر را اثبات می‌کنیم:

لم ۶-۳ فرض کنید $\{\eta_i\}_{i\in\mathbb{N}}$ دنباله‌ای از متغیرهای تصادفی *i.i.d.* با توزیع احتمال نمایی استاندارد باشد و θ,β اعداد مثبت دلخواهی باشند. داریم:

$$۰\leq\mathscr{E}_{\{\eta_i\}_i}\left\{\frac{۱}{\big(۱+\frac{۱}{\theta}\sum_{i=۱}^k\eta_i\big)^\beta}\right\}<۰/۸۳^k+\Big(\frac{۲\theta}{k}\Big)^\beta \qquad (۲۷-۶)$$

اثبات: از آن‌جا که η_iها و θ نامنفی‌اند، نامساوی سمت چپ بدیهی است. برای سادگی در ادامه اثبات تعریف می‌کنیم $\Gamma_k = \sum_{i=۱}^{k} \eta_i$. در نتیجه:

$$\begin{aligned} \mathscr{E}_{\{\eta_i\}}\Big\{\frac{۱}{(۱+\frac{\Gamma_k}{\theta})^{\beta}}\Big\} &= \int_{۰}^{\infty} \frac{۱}{(۱+\frac{\gamma}{\theta})^{\beta}} f_{\Gamma_k}(\gamma)d\gamma = \int_{۰}^{\frac{k}{۲}} \frac{۱}{(۱+\frac{\gamma}{\theta})^{\beta}} f_{\Gamma_k}(\gamma)d\gamma + \int_{\frac{k}{۲}}^{\infty} \frac{۱}{(۱+\frac{\gamma}{\theta})^{\beta}} f_{\Gamma_k}(\gamma)d\gamma \\ &\le \int_{۰}^{\frac{k}{۲}} f_{\Gamma_k}(\gamma)d\gamma + \int_{\frac{k}{۲}}^{\infty} \Big(\frac{۲\theta}{k}\Big)^{\beta} f_{\Gamma_k}(\gamma)d\gamma \le P\Big(\Gamma_k \le \frac{k}{۲}\Big) + \Big(\frac{۲\theta}{k}\Big)^{\beta} \end{aligned} \qquad (۲۸-۶)$$

همانطور که اشاره شد، Γ_k مجموع k متغیرتصادفی مستقل با میانگین واحد است که ایجاب می‌کند $\mathscr{E}\{\Gamma_k\} = k$. بنابراین $P(\Gamma_k \le \frac{k}{۲})$ یک احتمال انتهایی محسوب می‌شود. روش زیر راهی معروف برای نشان دادن میرایی نمایی این‌گونه احتمالات است:

$$\begin{aligned} P\Big(\Gamma_k \le \frac{k}{۲}\Big) &= P\Big(-\sum_{i=۱}^{k}\eta_i \ge -\frac{k}{۲}\Big) = P\Big(e^{-\sum_{i=۱}^{k}\eta_i} \ge e^{-\frac{k}{۴}}\Big) \le \frac{\mathscr{E}_{\{\eta_i\}_i}\{e^{-\sum_{i=۱}^{k}\eta_i}\}}{e^{-\frac{k}{۴}}} \\ &= e^{\frac{k}{۴}}\big(\mathscr{E}_{\eta}\{e^{-\eta}\}\big)^k = \Big(\frac{\sqrt{e}}{۲}\Big)^k < ۰/۸۳^k \end{aligned} \qquad (۲۹-۶)$$

نامساوی اول به کمک نامساوی کرافت معتبر است. ■

حال به سراغ $\mathscr{E}\{A_k\}$ می‌رویم. طبق تعریف داریم:

$$\mathscr{E}\{A_k\} = k^{۲-\alpha}\mathscr{E}_{\{\eta_i\}_i}\Big\{\big(\frac{\Gamma_۱}{\Gamma_{k+۱}}\big)^{\frac{۱}{\alpha}}\Big\} = k^{۲-\alpha}\mathscr{E}_{\eta_۱}\Big\{\mathscr{E}_{\{\eta_i\}_{i>۱}}\Big\{(۱+\frac{۱}{\eta_۱}\sum_{i=۲}^{k+۱}\eta_i)^{-\frac{۱}{\alpha}}\Big\}\Big\} \qquad (۳۰-۶)$$

اکنون با استفاده از لم ۶-۳ کران بالایی زیر حاصل می‌شود:

$$\mathscr{E}_{\{\eta_i\}_{i>۱}}\Big\{(۱+\frac{۱}{\eta_۱}\sum_{i=۲}^{k+۱}\eta_i)^{-\frac{۱}{\alpha}}\Big\} < ۰/۸۳^k + \big(\frac{۲\eta_۱}{k}\big)^{\frac{۱}{\alpha}} \qquad (۳۱-۶)$$

که در نهایت منجر به رابطه زیر می‌شود:

$$\mathscr{E}\{A_k\} \le k^{۲-\alpha}۰/۸۳^k + k^{۲-\alpha-\frac{۱}{\alpha}}\mathscr{E}_{\eta}\Big\{(۲\eta)^{\frac{۱}{\alpha}}\Big\} \qquad (۳۲-۶)$$

با توجه به $\alpha + \frac{۱}{\alpha} > ۲$ واضح است که $\lim_{k\to\infty}\mathscr{E}\{A_k\} = ۰$. مشابهاً برای بدست آوردن حد $\mathscr{E}\{B_{k,n}\}$ هنگامی که $k \to \infty$، به بیان لم زیر می‌پردازیم:

لم ۶-۴ فرض کنید $\{\eta_i\}_{i\in\mathbb{N}}$ دنباله‌ای از متغیرهای تصادفی *i.i.d.* با توزیع احتمال نمایی استاندارد باشد و $۱ < \beta$, $۰ < \theta$ اعداد دلخواهی باشند. داریم:

$$\mathscr{E}_{\{\eta_i\}_i}\Big\{۱+\sum_{k=۱}^{\infty}\frac{۱}{(۱+\frac{۱}{\theta}\sum_{i=۱}^{k}\eta_i)^{\beta}}\Big\} = ۱+\frac{\theta}{\beta-۱} \qquad (۳۳-۶)$$

اثبات: مشابه قبل، فرض کنید $\Gamma_k = \sum_{i=۱}^{k} \eta_i$. به دلیل i.i.d. بودن متغیرهای تصادفی، توزیع احتمال Γ_k از k مرتبه کانولوشن توزیع نمایی با خودش حاصل می‌شود:

$$f_{\Gamma_k}(\gamma) = \big(\underbrace{f_\eta * \cdots * f_\eta}_{k-\text{times}}\big)(\gamma) = \mathcal{F}_\omega^{-۱}\Big\{(۱+j\omega)^{-k}\Big\}(\gamma) \qquad (۳۴\text{-}۶)$$

در نتیجه با استفاده از نمایش حوزه فرکانس داریم:

$$\mathscr{E}_{\{\eta_i\}}\Big\{\frac{۱}{\big(۱+\frac{\Gamma_k}{\theta}\big)^{\beta}}\Big\} = \int_{۰}^{\infty}\int_{\mathbb{R}} \frac{۱}{\big(۱+\frac{\gamma}{\theta}\big)^{\beta}} \frac{e^{j\omega\gamma}}{(۱+j\omega)^k}\frac{d\omega}{۲\pi}\, d\gamma \qquad (۳۵\text{-}۶)$$

که نتیجه می‌دهد:

$$\begin{aligned}
\mathscr{E}_{\{\eta_i\}_i}\Big\{۱+\sum_{k=۱}^{\infty}\frac{۱}{\big(۱+\frac{\Gamma_k}{\theta}\big)^{\beta}}\Big\} &= \int_{۰}^{\infty}\int_{\mathbb{R}} \frac{e^{j\omega\gamma}}{\big(۱+\frac{\gamma}{\theta}\big)^{\beta}} \sum_{k=۰}^{\infty}\frac{۱}{(۱+j\omega)^k}\frac{d\omega}{۲\pi}\, d\gamma \\
&= \int_{۰}^{\infty}\int_{\mathbb{R}} \frac{e^{j\omega\gamma}}{\big(۱+\frac{\gamma}{\theta}\big)^{\beta}} \big(۱+\frac{۱}{j\omega}\big)\frac{d\omega}{۲\pi}\, d\gamma \\
&= \int_{\mathbb{R}} \frac{\delta(\gamma)+۱}{\big(۱+\frac{\gamma}{\theta}\big)^{\beta}} d\gamma = ۱ + \frac{\theta}{\beta-۱} < \infty \ \blacksquare
\end{aligned} \qquad (۳۶\text{-}۶)$$

نکته جالب توجه در لم فوق ارتباط آن با تابع زتای ریمان[۹] است: در لم ۶-۴ هنگامی که $\theta = ۱$، ما به دنبال یافتن امید ریاضی عبارتی به صورت $۱ + \sum_{k=۱}^{\infty} \Lambda_k^{-\beta}$ هستیم که $\{\Lambda_k\}_k$ دنباله‌ای از متغیرهای تصادفی با شرط $\Lambda_k \le \Lambda_{k+۱}$ و $\mathscr{E}\{\Lambda_k\} = k+۱$ است. در حالت حدی، هنگامی که واریانس این متغیرهای تصادفی به سمت صفر میل کند (عبور از حالت تصادفی به یقینی) این امید ریاضی به سمت $۱ + \sum_{k=۲}^{\infty} k^{-\beta}$ میل می‌کند که همان تابع زتای ریمان در نقطه β است. بنابراین امید ریاضی مورد نظر را می‌توان در حالت کلی به صورت تعمیم تصادفی تابع ریمان تفسیر کرد. نکته جالب توجه این است که $\beta = ۱$ در هر دو حالت یک قطب ساده است.

اکنون به بررسی حد $\mathscr{E}\{B_{k,n}\}$ بر می‌گردیم:

$$\mathscr{E}\{B_{k,n}\} = k^{\alpha-۲}\mathscr{E}\Big\{۱+\sum_{t=k+۲}^{n}\big(\frac{\Gamma_{k+۱}}{\Gamma_t}\big)^{\frac{۱}{\alpha}}\Big\} \le k^{\alpha-۲}\mathscr{E}\Big\{۱+\sum_{t=k+۲}^{\infty}\big(\frac{\Gamma_{k+۱}}{\Gamma_t}\big)^{\frac{۱}{\alpha}}\Big\} \qquad (۳۷\text{-}۶)$$

با استفاده از لم ۶-۴ داریم:

$$\begin{aligned}
\mathscr{E}_{\{\eta_i\}_{i>k+۱}}\Big\{۱+\sum_{t=k+۲}^{\infty}\big(\frac{\Gamma_{k+۱}}{\Gamma_t}\big)^{\frac{۱}{\alpha}}\Big\} &= \mathscr{E}_{\{\eta_i\}_{i>k+۱}}\Big\{۱+\sum_{t=k+۲}^{\infty}\frac{۱}{\Big(۱+\frac{\sum_{i=k+۲}^{t}\eta_i}{\Gamma_{k+۱}}\Big)^{\frac{۱}{\alpha}}}\Big\} \\
&= ۱ + \frac{\alpha}{۱-\alpha}\Gamma_{k+۱}
\end{aligned} \qquad (۳۸\text{-}۶)$$

[۹] Reiman Zeta Function

جدول ۶-۱: تعدادی از توزیع‌های فشرده‌پذیر با میرایی مرتبه $|t|^{-(q+1)}$

$f_x(t)$	Distribution
$\frac{q}{2\lambda}\left(1+\frac{\lvert t\rvert}{\lambda}\right)^{-(q+1)}$	Generalized Pareto
$\frac{\Gamma((q+1)/2)}{\Gamma(q/2)\sqrt{2\pi\lambda}}\left(1+\frac{t^2}{2\lambda}\right)^{-(q+1)/2}$	Student's t
$\frac{q}{2\lambda}e^{-(1+\frac{\lvert t\rvert}{\lambda})^{-q}}\left(1+\frac{\lvert t\rvert}{\lambda}\right)^{-(q+1)}$	Extreme Value
$\frac{(q/\lambda)(t/\lambda)^{q-1}}{\left(1+(t/\lambda)^q\right)^2}$	Log-Logistic

بنابراین:

$$\mathscr{E}\{B_{k,n}\} \quad \leq \quad k^{\alpha-2}\mathscr{E}_{\Gamma_{k+1}}\left\{1+\frac{\alpha}{1-\alpha}\Gamma_{k+1}\right\} = k^{\alpha-2}\left(1+\frac{\alpha(k+1)}{1-\alpha}\right) \qquad (۶-۳۹)$$

مجدداً از آن‌جا که $\alpha < 1$، کران بالای فوق در حالت حدی $k \to \infty$ به سمت صفر میل می‌کند. با استفاده از (۶-۲۶) و این موضوع که $\mathscr{E}\{A_k\} = 0$ و $\mathscr{E}\{B_{k,n}\} = 0$ در حالت حدی صفر هستند، می‌دانیم که:

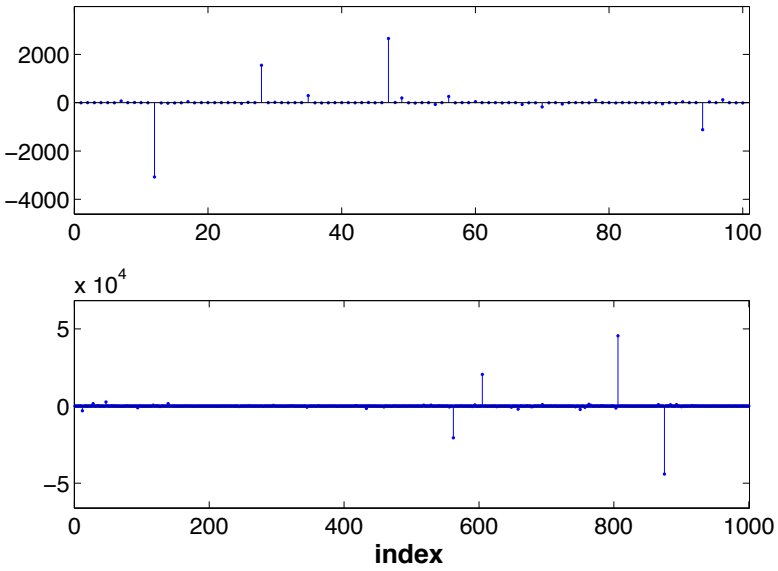
$$\lim_{k\to\infty} P\left(\frac{\Gamma_{k+1}^{-\frac{1}{\alpha}} + \cdots + \Gamma_n^{-\frac{1}{\alpha}}}{\Gamma_1^{-\frac{1}{\alpha}} + \cdots + \Gamma_n^{-\frac{1}{\alpha}}} \leq 1-r\right) = 1 \qquad (۶-۴۰)$$

دقت کنید که کران بالای اثبات شده برای $\mathscr{E}\{A_k\}$ و $\mathscr{E}\{B_{k,n}\}$ در (۶-۳۲) و (۶-۳۹)، تنها تابعی از k و مستقل از n هستند. پس برای این که احتمال طرح شده در (۶-۴۰) برای $0 < \epsilon < 1$ دلخواه از $\frac{\epsilon}{2} - 1$ بزرگتر باشد، تنها کافی است k از حدی بزرگتر اختیار شود ($k > k_0$). از طرف دیگر، قضیه ۶-۲ نشان می‌دهد که برای این مقدار k_0 (که تنها به ϵ بستگی دارد)، n به اندازه کافی بزرگ، اختلاف بین دو احتمال در (۶-۲۴) از $\frac{\epsilon}{2}$ کوچکتر خواهد بود. بنابراین برای این مقدار k_0 و nهای به اندازه کافی بزرگ داریم:

$$P\left(\frac{\sigma_1\left(k_0;\{x_i\}_{i=1}^n\right)}{\sigma_1\left(n;\{x_i\}_{i=1}^n\right)} \geq r\right) \geq 1-\epsilon \qquad (۶-۴۱)$$

به وضوح هنگامی که n رشد می‌کند، $\frac{k_0}{n} \to 0$، که اثبات را به پایان می‌برد. ■

قضیه ۶-۳ تقریباً بیان می‌کند که توزیع احتمال $f_x(t)$ با میرایی چندجمله‌ای از مرتبه $|t|^{-(\alpha+1)}$ که $\alpha < 1$، ℓ_1-فشرده‌پذیر است. حال فرض کنید میرایی توزیع احتمال از مرتبه $|t|^{-(q+1)}$ باشد که q لزوماً کمتر از یک نیست. در این صورت توزیع احتمال $y = |x|^p$ که برابر است با $f_y(t) = \frac{f_x(t^{1/p})+f_x(-t^{1/p})}{pt^{1-1/p}}$ از مرتبه $|t|^{-(q/p+1)}$ میرا خواهد بود. لذا با توجه به لم ۶-۲، f_x برای هر $p > q$، ℓ_p-فشرده‌پذیر خواهد بود. جدول ۶-۱ تعدادی از



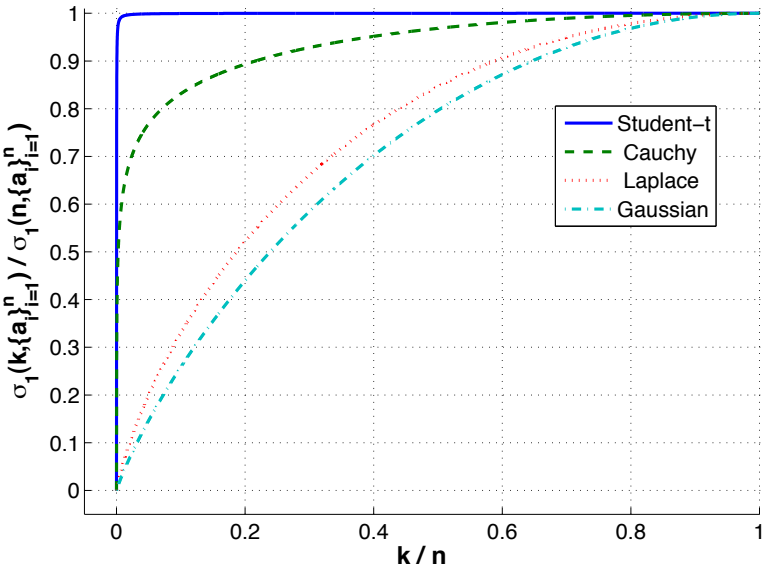
شکل ۶–۱: یک تحقق زمانی از دنباله تصادفی و .i.i.d با توزیع احتمال Student's t و پارامتر $q = ۰/۵$.

توزیع احتمال‌های میرا از مرتبه $|t|^{-(q+1)}$ را نشان می‌دهد که مطابق آن چه گفته شد، فشرده‌پذیر نیز هستند. جالب آن که این توزیع‌ها با استدلال‌های مطرح شده در [۲۷] مطابقت دارند. علاوه بر این توزیع‌ها، تمامی توزیع‌های α–پایدار برای $۰ < \alpha < ۲$ که شامل توزیع کوشی و Lévy نیز می‌شود، برای $p > \alpha$، ℓ_p–فشرده‌پذیر هستند.

۶–۵ نتایج عددی

تعریف فشرده‌پذیری در دنباله‌های تصادفی بیان می‌کند که تحقق‌های این دنباله‌ها با احتمال بالا فشرده‌پذیرند. برای بررسی این مطلب، در شکل ۶–۱ یک تحقق از دنباله .i.i.d با توزیع Student's t با پارامتر $q = ۰/۵$ (میرایی از مرتبه $|t|^{-۰/۵}$) رسم شده است. در این شکل یک بار ۱۰۰ جمله اول و بار دیگر ۱۰۰۰ جمله اول که شامل ۱۰۰ جمله قبلی هم می‌شوند، دیده می‌شود. در هر دو شکل مشاهده می‌شود که تعداد جملاتی که مقدار قابل توجهی دارند، اندک است. اما نکته جالب‌تر از آن این است که تمام جملات پر اهمیت در ۱۰۰ جمله اول، جزء جملات بسیار کوچک در پنجره ۱۰۰۰ تایی بشمار می‌آیند. این امر نشان می‌دهد که با جلو رفتن در دنباله‌های .i.i.d فشرده‌پذیر، همواره شاهد رشد جملات اصلی هستیم به نحوی که باعث کم اهمیت شدن جملات قبلی می‌شوند.

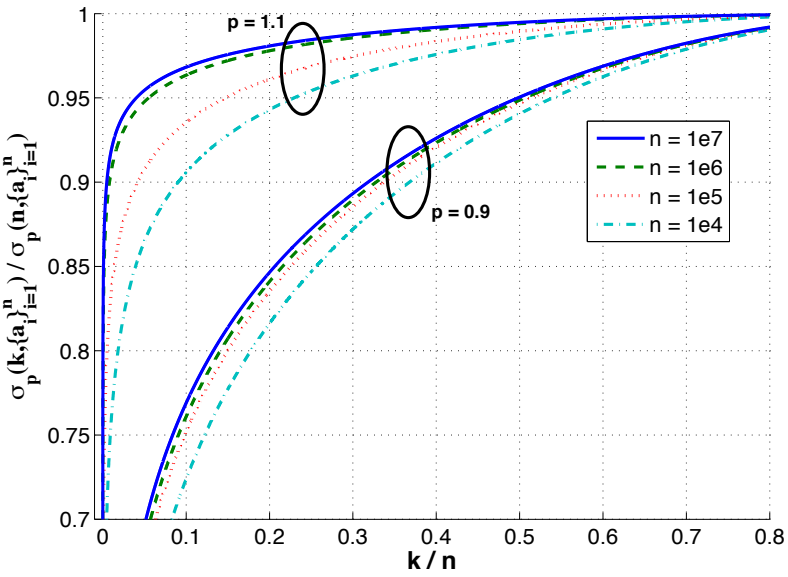
برای به تصویر کشیدن مفهوم فشرده‌پذیری در دنباله‌های تصادفی، مقادیر متوسط $\frac{\sigma_۱(k;\{a_i\}_{i=۱}^n)}{\sigma_۱(n;\{a_i\}_{i=۱}^n)}$ در ۵۰۰ تحقق متفاوت را بر حسب k و برای توزیع احتمال‌های مختلف در شکل ۶–۲ رسم کرده‌ایم. برای این



شکل ۶-۲: نمودارهای متوسط تابع $\frac{\sigma_p(k;\{a_i\}_{i=1}^n)}{\sigma_p(n;\{a_i\}_{i=1}^n)}$ بر حسب $\frac{k}{n}$ برای تحقق‌های i.i.d. از توزیع احتمال‌های گوسی، لاپلاس، کوشی و Student's t ($q = ۰/۵$) به ازای $p = ۱$ و $n = ۱۰^۴$.

نتایج، طول دنباله $n = ۱۰^۴$ قرارداده شده است. قضیه ۶-۳ بیان می‌کند که توزیع Student's t با $q = ۰/۵$ از نوع $\ell_۱$-فشرده‌پذیر است و همان‌گونه که در شکل ۶-۲ دیده می‌شود، نمودار در مقادیر خیلی کوچک $\frac{k}{n}$ اشباع می‌شود. فشرده‌پذیری دنباله ایجاب می‌کند که در nهای بزرگ این نمودار به سمت تابع پله میل می‌کند. در نقطه مقابل، توزیع‌های گوسی و لاپلاس فشرده ناپذیرند و در نتیجه نمودار حدی آن‌ها هنگامی که $n \to \infty$ فرق چندانی با نمودارهای نمایش داده شده در شکل ۶-۲ ندارند. در این مثال توزیع کوشی در مرز $\ell_۱$-فشرده‌پذیری قرار دارد؛ از آن‌جا که میرایی این توزیع از مرتبه $t^{-۲}$ است، برای هر $p > ۱$، $\ell_۱$-فشرده‌پذیر است اما در مورد $p = ۱$ با استفاده از قضایای مطرح شده، نتیجه‌ای نمی‌توان گرفت.

برای بررسی بیشتر اثر p در ℓ_p-فشرده‌پذیری، مقدار متوسط کسر انرژی را برای توزیع کوشی با دو نرم $\ell_{۰/۹}$ و $\ell_{۱/۱}$ در شکل ۶-۳ رسم کرده‌ایم. همانطور که قضیه ۶-۳ پیش‌بینی می‌کند، نمودارهای مربوط به نرم $\ell_{۱/۱}$ با افزایش n به سمت تابع پله میل می‌کنند ولی در حالت $\ell_{۰/۹}$ چنین به نظر نمی‌رسد.



شکل ۶-۳: نمودارهای متوسط تابع $\frac{\sigma_p(k;\{a_i\}_{i=1}^n)}{\sigma_p(n;\{a_i\}_{i=1}^n)}$ بر حسب $\frac{k}{n}$ در nهای مختلف برای تحقق‌های .i.i.d از توزیع احتمال کوشی هنگامی که $p = ۰/۹$ و $p = ۱/۱$ مورد نظر باشند.

فصل ۷

کاهش رتبه در ماتریس‌ها به‌کمک عملگرهای المانی

۷-۱ مقدمه

یکی از مسائل مهم در مورد سیگنال‌های واقعی، درجه آزادی موجود در آن‌هاست. در مبحث کاهش بعد[۱] تمرکز اصلی بر روی یافتن تعداد متغیرهای مستقل در ساختار یک گروه از سیگنال‌هاست که به فشرده‌سازی اطلاعات منجر می‌شود. یافتن بعد ذاتی در بسیاری در زمینه‌ها همچون یادگیری ماشین[۲] [۶۱]، بینایی ماشین[۳] [۹۲]، شبکه حسگرها[۴] [۷۶]، کاهش پهنای باند[۵] [۷۲] و فشرده‌سازی اطلاعات[۶] [۲۰] اهمیت بسزایی دارد. دلیل اصلی کاهش بعد، نمایش فشرده داده بدون از دست دادن اطلاعات است. ساده‌سازی در ساختار، ما را قادر به استفاده از الگوریتم‌های طراحی شده برای بعد پایین (که نسبتاً سریع هستند) در مورد سیگنال‌های با بعد بالا می‌سازد. یکی از روش‌های رایج در نمایش داده‌ها، استفاده از ماتریس است. فرض کنید $\mathbf{B}$ به نوعی بیانگر اطلاعات اصلی باشد؛ در مبحث کاهش بعد به دنبال یافتن ماتریس $\mathbf{A}$ای با مرتبه کم هستیم که به بهترین نحو ماتریس $\mathbf{B}$ را تقریب بزند. در حقیقت، رتبه یک ماتریس بیانگر درجه آزادی آن است. یکی از نتایج معروف در این زمینه آن است که بهترین تقریب رتبه k از دیدگاه اندازه فروبنیوسی[۷] برای ماتریس $\mathbf{B}$، از نگه داشتن بزرگترین k مقدار

[۱] Dimensionality Reduction
[۲] Machine Learning
[۳] Computer Vision
[۴] Sensor Networks
[۵] Bandwidth Reduction
[۶] Data Compression
[۷] Frobenius Norm

تکین[۸] ماتریس و صفر کردن بقیه مقادیر تکین در تجزیه SVD ($\mathbf{b}_k$) بدست می‌آید. به عبارت بهتر، می‌توان نشان داد برای هر ماتریس $\mathbf{M}$ با مرتبه k داریم :

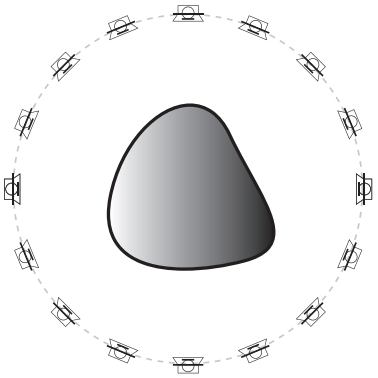
$$|\mathbf{B} - \mathbf{B}_k|_F \leq |\mathbf{M} - \mathbf{B}|_F$$

که در آن $|\cdot|_F$ بیانگر اندازه فروبنیوسی است. این روش موفقیت چشم‌گیری در زمینه‌هایی همچون بازیابی اطلاعات[۹] [۱۳] ، تشخیص چهره[۱۰] [۹۲] و تکمیل ماتریس[۱۱] [۲۱، ۶۴] که در آن‌ها ماتریس $\mathbf{B}$ تقریباً کم مرتبه است، بدست آورده است. در این فصل، کاهش بعد از دیدگاه پردازش سیگنال مورد بررسی قرار می‌گیرد که در مباحثی همچون شبکه حسگرهای بی‌سیم و توموگرافی فراصوتی، کاربرد دارد. در این موارد، سیستم شامل چندین گره ارتباطی است به گونه‌ای که کانال ارتباطی باعث اعوجاج بر روی پیام‌های ارسالی می‌شود. در بسیاری از مواقع اگر پیام‌های ارسالی بین زوج مرتب‌های گیره را به صورت ماتریس $\mathbf{A}$ نمایش دهیم، ماتریس بدست آمده رتبه پایینی دارد اما به دلیل اعوجاج تولید شده توسط کانال ارتباطی، ماتریس حاصل از سیگنال‌های دریافتی ($\mathbf{B}$) دچار افزایش رتبه شده است. نکته قابل توجه در این است که اعوجاج صورت گرفته توسط کانال بر روی تمام درایه‌های ماتریس یکسان است. بطور مثال اگر تاثیر کانال معادل با جذر گرفتن باشد، تمام سیگنال‌های دریافتی بطور یکسان تحت تاثیر عملگر جذر قرار می‌گیرند. از این رو، با وجود آن که تعداد پارامترهای مستقل در تولید ماتریس $\mathbf{B}$ به همان اندازه ماتریس $\mathbf{A}$ است، به دلیل اعوجاج غیرخطی کانال، رتبه $\mathbf{B}$ نمی‌تواند به درستی بیانگر این درجه آزادی باشد. حال آن‌که اگر بتوان عکس اعوجاج ایجاد شده توسط کانال را بر روی درایه‌ها عمل کرد، مجدداً به ماتریس مرتبه کم $\mathbf{A}$ با نمایش صحیح درجه آزادی می‌رسیم. در ادامه به طور مختصر به شرح دو مثال ذکر شده، یعنی توموگرافی فراصوتی و شبکه‌ی حسگرها می‌پردازیم.

کالیبره کردن در توموگرافی فراصوتی

در توموگرافی فراصوتی، مجموعه حسگرها معمولاً بر روی یک حلقه محیط بر جسم نامعلوم قرار دارند که هریک از این حسگرها می‌تواند در نقش گیرنده یا فرستنده عمل کند. در هر لحظه، تنها یکی از حسگرها به عنوان فرستنده، سیگنالی ارسال می‌کند که پس از برخورد با جسم نامعلوم توسط مابقی حسگرها ضبط می‌شود.

[۸] Singular Value
[۹] Information Retrieval
[۱۰] Face Recognition
[۱۱] Matrix Completion



شکل ۷–۱: آرایش دایروی حسگرها در توموگرافی فراصوتی

این عمل به طور مرتب و چرخشی توسط تمامی حسگرها صورت می‌گیرد و در نهایت براساس تمام بازتاب‌های دریافتی، خواصی از قبیل شکل و ضریب سختی جسم نامعلوم بدست می‌آید. ساختار دو بعدی فوق را می‌توان به طور مشابهی به سه بعد تعمیم داد. شکل ۷–۱ ساختار دو بعدی را نشان می‌دهد.

برای پردازش بازتاب‌های بدست آمده، ابتدا لازم است که فاصله دقیق میان هر دو زوج حسگرها محاسبه شود. برای این منظور، عملیات ارسال و دریافت ابتدا بدون قراردادن جسم در داخل محفظه انجام می‌شود تا تاخیرهای زمانی میان هر ارسال و دریافت به صورت t_{ij} بدست آید. ماتریس $\mathbf{T} = [t_{ij}]$ بیانگر کلیه تاخیرها و در نتیجه فواصل میان زوج حسگرهاست. در حالت کلی، ماتریس $\mathbf{T}$ دارای رتبه کامل است اما در [۶۲] نشان داده شده است که مرتبه ماتریس $\tilde{\mathbf{T}} = [t_{i,j}^{۲}]_{n\times n}$ حداکثر ۳ است. از آن‌جا که مشاهدات معمولاً همراه با نویز است، رتبه این ماتریس دقیقاً ۳ نخواهد بود و با پیدا کردن بهترین تقریب رتبه ۳ آن، نه تنها نویز کاهش می‌یابد بلکه نمایش فشرده‌تری از این ماتریس بدست می‌آید. این مثال، حالتی را نشان می‌دهد که در آن مرتبه ذاتی ۳ است و تاثیر اعوجاج کانال، معادل جذر گرفتن است. در نتیجه عملگر المانی لازم برای کاهش بعد، به توان ۲ رساندن است. پدیده مشابهی در تخمین زمان دریافت سیگنال‌های صوتی توسط یک آرایه از میکروفون‌ها رخ می‌دهد [۳۰].

مکان یابی حسگرها در شبکه حسگرهای بی‌سیم

در مخابرات رادیویی، توان سیگنال دریافتی به صورت تابعی از فاصله میان فرستنده و گیرنده افت می‌کند که به پدیده افت مسیر[۱۲] شهرت دارد. در حالت کلی اگر فاصله فرستنده و گیرنده r باشد، توان دریافتی متناسب با $\frac{۱}{r^{\alpha}}$ افت می‌کند که توان α بین ۲ (در فضای آزاد) تا ۶ (در فضای پرانعکاس) تغییر می‌کند [۷۹].

در مسأله‌ی مکان یابی حسگرها، n حسگر در فضای d بعدی پخش شده‌اند و هر حسگر فاصله‌ی خود از سایر حسگرها (احتمالاً حسگرهای مجاور) را اندازه می‌گیرد. در عمل، اندازه‌گیری فاصله براساس توان دریافتی صورت می‌پذیرد و فرض کنید $d_{i,j}$ بیانگر فاصله اقلیدسی میان حسگرهای i و j، و $\mathbf{D} = [d_{i,j}]$ ماتریس فاصله باشد، همچنین فرض کنید $\mathbf{P} = [p_{i,j}]_{n\times n}$ ماتریس توان‌های دریافتی باشد که در آن $p_{i,j}$ توان دریافتی حسگر iام از سیگنال ارسالی حسگر jام است. واضح است که با داشتن $\mathbf{D}$ و تعدادی حسگر مرجع، موقعیت تمامی حسگرها مشخص خواهد شد اما در عمل به جای ماتریس $\mathbf{D}$، ماتریس $\mathbf{P}$ در دست است. با وجود آن که $\mathbf{P}$ در حالت کلی رتبه کامل دارد، به راحتی می‌توان نشان داد که ماتریس $\tilde{\mathbf{P}} = [p_{i,j}^{-\frac{۲}{\alpha}}]_{n\times n}$ حداکثر رتبه $d^{۲}$ دارد. در نتیجه، تخمین ماتریس $\tilde{\mathbf{P}}$ و استفاده از خاصیت کم‌رتبه بودن آن، کمک بسزایی در مکان یابی حسگرها خواهد داشت [۶۲، ۴۲]. از آن‌جا که α یکی از خواص محیط محسوب می‌شود، معمولاً مقدار آن در دست نیست، در نتیجه این وضعیت با آن که بعد ذاتی مسئله $d^{۲}$ است، به دلیل اعمال یک عملگر المانی نامشخص، رتبه ماتریس کامل شده است.

در این فصل، به بررسی مسئله ذکر شده در حالت کلی می‌پردازیم که در آن یک ماتریس کم‌رتبه توسط یک عملگر المانی شبه چندجمله‌ای و نامعلوم به یک ماتریس مرتبه کامل تبدیل شده است و با در دست داشتن ماتریس رتبه کامل، به دنبال یافتن عملکرد اعمال شده و در نتیجه ماتریس کم‌رتبه هستیم [۲]. در ادامه، پس از معرفی تعدادی نماد، مسئله را به طور دقیق و در قالب روابط ریاضی بیان می‌کنیم. سپس مسئله را در دو حالت توان صحیح و توان حقیقی، مورد بررسی قرار می‌دهیم.

۷-۲ نمادها و شرح مساله

در این فصل از نماد $\odot$ برای بیان عملگر المانی بر روی درایه‌های یک ماتریس استفاده می‌کنیم. به عنوان مثال اگر $\mathbf{A} = [a_{i,j}]_{m\times n}$، در این صورت $\mathbf{A}^{\odot ۲} = [a_{i,j}^{۲}]_{m\times n}$ و $\odot f(\mathbf{A}) = \big[f(a_{i,j})\big]_{m\times n}$. ماتریس رتبه‌ناقص و رتبه آن

[۱۲] Path-loss

به ترتیب با $\mathbf{A}$ و k نمایش داده می‌شود و اگر $\mathbf{B} = \odot f(\mathbf{A})$، $f(.)$ عملگر اعوجاجی رابط بین $\mathbf{A}$ و $\mathbf{B}$ نامیده می‌شود. همچنین وارون f در صورت وجود، عملگر کاهنده رتبه تلقی می‌شود. به علاوه نمونه نویزی شده $\mathbf{B}$ را با $\tilde{\mathbf{B}}$ نشان می‌دهیم. به دلیل استفاده مکرر از دترمینان دسته‌ای از ماتریس‌ها، نمادهای زیر را تعریف می‌کنیم:

$$T_{\mathbf{B}}(x) = \det\left(\mathbf{B}^{\odot x}\right) \quad (۱-۷)$$

$$\bar{T}_{\mathbf{B}}(q_1, \ldots, q_n) = \det\left(\left[(\ln b_{i,j})^{q_i}\right]\right) \quad (۲-۷)$$

شرح مسئله : فرض کنید $f(x) = x^p$ عملگر اعوجاجی باشد که در آن $p > 0$ نامشخص است. همچنین فرض کنید $\mathbf{A}$ ماتریس رتبه‌ناقص باشد و $\mathbf{B} = \odot f(\mathbf{A})$. اکنون هدف تخمین عملگر f و در نتیجه ماتریس رتبه‌ناقص $\mathbf{A}$ با در دست داشتن $\mathbf{B}$ و یا نمونه‌ی نویزی آن ($\tilde{\mathbf{B}}$) است .

۷-۳ توان صحیح

قضیه زیر اثر یک عملگر اعوجاجی از نوع چندجمله‌ای بر روی یک ماتریس رتبه‌ناقص را نشان می‌دهد.

قضیه ۷-۱ فرض کنید رتبه ماتریس $\mathbf{A}_{n\times n}$، k باشد و p یک عدد طبیعی دلخواه باشد. در این صورت داریم :

$$\operatorname{rank}\left(\mathbf{A}^{\odot p}\right) \le \min\left\{n, \binom{k+p-1}{p}\right\} \quad (۳-۷)$$

اثبات: از آن‌جا که $\operatorname{rank}(\mathbf{A}) = k$، می‌توانیم k سطر مستقل خطی $\{\mathbf{v}_i\}_{i=1}^{k}$ را از بین سطرهای $\mathbf{A}$ انتخاب کنیم. در نتیجه، تمام سطرهای $\mathbf{A}$ را می‌توان بر حسب ترکیب خطی این سطرها نمایش داد:

$$\mathbf{A}_{n\times n} = \underbrace{\begin{bmatrix} c_{1,1} & \cdots & c_{1,k} \\ \vdots & \ddots & \vdots \\ c_{n,1} & \cdots & c_{n,k} \end{bmatrix}}_{\mathbf{C}_{n\times k}} \underbrace{\begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_k \end{bmatrix}}_{\mathbf{V}_{k\times n}} = \left[\sum_{l=1}^{k} c_{i,l} v_{l,j}\right] \quad (۴-۷)$$

از این رو داریم:

$$\begin{aligned} \mathbf{A}^{\odot p} &= \left[\left(\sum_{l=1}^{k} c_{i,l} v_{l,j}\right)^p\right] = \left[\sum_{p_1+\cdots+p_k=p} \binom{p}{p_1, \ldots, p_k} \prod_{l=1}^{k} (c_{i,l} v_{l,j})^{p_l}\right] \\ &= \sum_{p_1+\cdots+p_k=p} \binom{p}{p_1, \ldots, p_k} \left[\prod_{l=1}^{k} (c_{i,l} v_{l,j})^{p_l}\right] \end{aligned} \quad (۵-۷)$$

که در آن $\left[\prod_{l=۱}^{k}(c_{i,l}v_{l,j})^{p_l}\right]$ نمایش دهنده یک ماتریس $n \times n$ با درایه‌ی i,j به صورت $\prod_{l=۱}^{k}(c_{i,l}v_{l,j})^{p_l}$ است.

پس:

$$\operatorname{rank}(\mathbf{A}^{\odot p}) \leq \sum_{p_۱+\cdots+p_k=p} \operatorname{rank}\left(\left[\prod_{l=۱}^{k}(c_{i,l}v_{l,j})^{p_l}\right]\right) \tag{۶–۷}$$

دقت کنید که :

$$\left[\prod_{l=۱}^{k}(c_{i,l}v_{l,j})^{p_l}\right] = \begin{bmatrix} \prod_{l=۱}^{k} c_{۱,l}^{p_l} \\ \vdots \\ \prod_{l=۱}^{k} c_{n,l}^{p_l} \end{bmatrix} \begin{bmatrix} \prod_{l=۱}^{k} v_{l,۱}^{p_l} & \cdots & \prod_{l=۱}^{k} v_{l,n}^{p_l} \end{bmatrix} \tag{۷–۷}$$

که نتیجه می‌دهد:

$$\operatorname{rank}\left(\left[\prod_{l=۱}^{k}(c_{i,l}v_{l,j})^{p_l}\right]\right) = ۱ \tag{۸–۷}$$

با ادغام این رابطه و رابطه‌ی (۷–۶)، بدست می‌آوریم:

$$\operatorname{rank}(\mathbf{A}^{\odot p}) \leq \sum_{p_۱+\cdots+p_k=p} ۱ = \binom{k+p-۱}{p} \tag{۹–۷}$$

که اثبات را کامل می‌کند. ■

نکته ۳ اگر $\mathbf{A}$ یک ماتریس چرخشی باشد به طوری که تبدیل DFT سطر اول تنها k درایه ناصفر و متوالی داشته باشد، در این صورت :

$$\operatorname{rank}(\mathbf{A}^{\odot p}) \leq p(k-۱)+۱ \tag{۱۰–۷}$$

دلیل این امر آن است که ضرب در حوزه زمان معادل با کانولوشن در حوزه فرکانس است و مشابهاً به توان p رساندن در حوزه زمان معادل با p مرتبه کانولوشن در حوزه فرکانس است. علاوه بر این، می‌دانیم که ماتریس‌های چرخشی را می‌توان به کمک ماتریس‌های یکه DFT و $IDFT$ به صورت قطری تجزیه کرد که در آن عناصر روی قطر، ضرایب DFT سطر اول هستند.

نکته ۴ عملگر اعوجاجی در نظر گرفته شده در قضیه ۷–۱، حالت خاصی از عملکرد چندجمله‌ای است که در حالت کلی به صورت $\mathbf{B} = \odot f(\mathbf{A})$ با شرط $f(x) = \sum_{i=۰}^{p} f_i x^i$ بیان می‌شود. در حقیقت، $f(x)$ در قضیه ۷–۱، یک جمله‌ای است. برای حالت کلی چندجمله‌ای داریم:

$$\odot f(\mathbf{A}) = \sum_{i=۰}^{p} f_i \mathbf{A}^{\odot i} \tag{۱۱–۷}$$

و در نتیجه:

$$\operatorname{rank}(\odot f(\mathbf{A})) \leq \sum_{i=0}^{p} \operatorname{rank}(\mathbf{A}^{\odot i}) \leq \sum_{i=0}^{p} \binom{k+i-1}{i} = \binom{k+p}{p} \qquad (۱۲-۷)$$

نکته ۵ کران مذکور در قضیه ۷-۱ اغلب حاصل می‌شود. این نکته در آشکارسازی عملگرهای اعوجاجی چندجمله‌ای کارساز است. فرض کنید $f(x)$ چندجمله‌ای درجه p و $\mathbf{A}_{n\times n}$ ماتریسی با رتبه $k_{\mathbf{A}}$ باشد که در آن $k_{\mathbf{A}}$ نسبت به n بسیار کوچک است. همچنین، $\mathbf{B} = \odot f(\mathbf{A})$ ماتریس اعوجاج یافته با رتبه $k_{\mathbf{B}} \leq \binom{k_{\mathbf{A}}+p}{p}$ است. اگر $\mathbf{B}$ یک ماتریس رتبه‌ناقص حالت کلی می‌بود، ماتریس $\mathbf{B}^{\odot i}$ $(i \in \mathbb{N})$ با احتمال زیاد، رتبه $\binom{k_{\mathbf{B}}+i-1}{i}$ می‌داشت اما از آن‌جا که $\mathbf{B}$ خود از یک ماتریس رتبه‌ناقص $\mathbf{A}$ حاصل شده است، $\mathbf{B}^{\odot i}$ توسط یک چندجمله‌ای از درجه pi به $\mathbf{A}$ مربوط می‌شود. در نتیجه رتبه آن نمی‌تواند از $\binom{k_{\mathbf{A}}+pi}{pi}$ فراتر رود. به سادگی می‌توان بررسی نمود که این کران، کمتر از کران پیش بینی شده توسط حالت کلی نکته فوق است. از این رو به راحتی می‌توان $\mathbf{B}$ را از یک ماتریس رتبه‌ناقص حالت کلی متمایز کرد. البته به شرطی که $k_{\mathbf{A}}$ به اندازه کافی از n کوچکتر باشد. در غیر این صورت $\mathbf{B}^{\odot i}$ و یا $\mathbf{B}$ احتمالاً ماتریس‌های رتبه کامل هستند. علاوه بر این، روند تغییرات $rank(\mathbf{B}^{\odot i})$ نسبت به i در یافتن درجه چندجمله‌ای اعوجاجی نقش موثری دارد.

۷-۴ توان حقیقی

در این بخش، تمرکز بر روی حالت $\mathbf{B} = \mathbf{A}^{\odot p}$ است که در آن p. یک تابع معکوس‌پذیر فرض می‌شود (که بتوانیم ماتریس $\mathbf{A}$ را بازیابی کنیم). به عنوان مثال، حالت $\mathbf{A}^{\odot \frac{3}{5}}$ را در نظر بگیرید؛ در صورتی که المان‌های $\mathbf{A}$ حقیقی باشند هر دوی $\frac{3}{5}$ و $\frac{5}{3}$. وارون‌پذیر و خوش تعریف هستند. برای ماتریس رتبه‌ناقص $\mathbf{A}$، ماتریس $\mathbf{B} = \mathbf{A}^{\odot \frac{3}{5}}$ به جز حالت‌های خاص، رتبه کامل است. در این‌جا با مشاهده $\mathbf{B}_{n\times n}$ می‌خواهیم به وجود و یا عدم وجود ماتریس رتبه‌ناقص $\mathbf{A}$ از پس $\mathbf{B}$ پی ببریم و در صورت وجود، عملگر p. را تخمین بزنیم. دقت کنید که $\mathbf{B}^{\odot \frac{1}{p}}$ همان ماتریس رتبه‌ناقص اصلی است اما اگر x یک تقریب بسیار خوب از $\frac{1}{p}$ باشد، $\mathbf{B}^{\odot x}$ کماکان دارای رتبه کامل است. به بیان بهتر، حتی تقریب‌های بسیار خوب از $\frac{1}{p}$، رتبه را کاهش نمی‌دهند. این مشکل اساساً به دلیل گسسته بودن مقادیر رتبه است؛ برای غلبه بر این مشکل باید معیاری پیوسته برای رتبه یک ماتریس معرفی کنیم. برای این منظور از تابع $T_{\mathbf{B}}(x)$ که در (۷-۱) تعریف شده است استفاده می‌کنیم. واضح است که اگر $\mathbf{B}$ شامل درایه ۰ نباشد، $\mathbf{B}^{\odot 0} = 1_{n\times n}$ است. در نتیجه،برای $n > 1$ همواره $T_{\mathbf{B}}(0) = 0$ همچنین اگر $\mathbf{B} = \mathbf{A}^{\odot p}$ که در آن

$\mathbf{A}$ ماتریس رتبه‌ناقص است، $T_{\mathbf{B}}(\frac{۱}{p}) = T_{\mathbf{A}}(۱) = ۰$. دقت کنید که $T_{\mathbf{B}}(x)$ تابعی پیوسته از x است و در نتیجه اگر x به اندازه کافی به $\frac{۱}{p}$ نزدیک باشد، $T_{\mathbf{B}}(x)$ نیز به اندازه کافی به ۰ نزدیک خواهد بود. بنابراین ریشه‌های $T_{\mathbf{B}}(x)$ به غیر از حالت بدیهی $x = ۰$، نقش کلیدی در آشکارسازی درجه آزادی به کار رفته در ساختار $\mathbf{B}$ ایفا می‌کند. با این وجود یافتن ریشه‌های $T_{\mathbf{B}}(x)$ کار آسانی نیست. به منظور یافتن ریشه‌های $T_{\mathbf{B}}(x)$ از تقریب تابع ذکر شده توسط بسط تیلور آن استفاده می‌کنیم.

لم ۷-۱ تابع $T_{\mathbf{B}}(x)$ در هر نقطه بسط تیلور همگرا دارد و جمله qام در بسط حول $x = ۰$ عبارت است از $(T_{\mathbf{B}}(x) = \sum_{q=۰}^{\infty} t_q x^q)$:

$$t_q = \frac{\sum_{\pi \in S_n} \operatorname{sgn}(\pi) \big(\sum_{i=۱}^{n} \ln b_{i,\pi(i)} \big)^q}{q!} \qquad (۷\text{-}۱۳)$$

که در آن S_n بیانگر مجموعه تمام جایگشت‌های $\{۱, \ldots, n\}$ است (که تعداد آن‌ها برابر با $n!$ است) و برای هر $\pi \in S_n$، علامت π (که با $\operatorname{sgn}(\pi)$ نمایش داده می‌شود) برابر با $(-۱)^{N(\pi)}$ است که $N(\pi)$ کمترین تعداد جابه‌جایی‌های دوتایی لازم برای تبدیل جایگشت π به جایگشت همانی را نشان می‌دهد.

اثبات: برای بدست آوردن ضرایب سری تیلور، از بسط دترمینان به صورت زیر استفاده می‌کنیم:

$$\begin{aligned} T_{\mathbf{B}}(x) &= \det\left[b_{i,j}^x\right] = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=۱}^{n} b_{i,\pi(i)}^x = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) e^{x \sum_{i=۱}^{n} \ln b_{i,\pi(i)}} \\ &= \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \sum_{q=۰}^{\infty} \frac{\big(\sum_{i=۱}^{n} \ln b_{i,\pi(i)} \big)^q}{q!} x^q = \sum_{q=۰}^{\infty} \frac{x^q}{q!} \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \big(\sum_{i=۱}^{n} \ln b_{i,\pi(i)} \big)^q \end{aligned} \qquad (۷\text{-}۱۴)$$

رابطه‌ی آخر، ضرایب سری تیلور را نشان می‌دهد. دقت کنید $\det(\mathbf{M}_۱ + \mathbf{M}_۲)$ لزوماً برابر با $\det(\mathbf{M}_۱) + \det(\mathbf{M}_۲)$ نیست و لذا $\frac{\partial^q}{\partial x^q} T_{\mathbf{B}}(x)$ لزوماً برابر با $\det\left(\frac{\partial^q}{\partial x^q} \mathbf{B}^{\odot x}\right)$ نیست که اهمیت نحوه بدست آوردن رابطه (۷-۱۴) را نشان می‌دهد. همچنین به دلیل نمایش $T_{\mathbf{B}}(x)$ به صورت جمع، تعداد متناهی جمله نمایی، سری تیلور مذکور رد تمامی نقاط همگراست. ■

با وجود آن که لم ۷-۱ ضرایب سری تیلور $T_{\mathbf{B}}(x)$ را بیان می‌کند، نمایش مفید دیگری از سری تیلور وجود دارد که در ادامه آن برای بدست آوردن یک کران بر روی خطای حاصل از تقریب $T_{\mathbf{B}}(x)$ توسط متناهی جمله از سری تیلور آن، استفاده می‌کنیم.

قضیه ۷–۲ $n-1$ جمله اول سری تیلور $T_{\mathbf{B}}(x)$ همگی صفر هستند و داریم :

$$T_{\mathbf{B}}(x) = \sum_{q=n-1}^{\infty} \frac{x^q}{q!} \sum_{\substack{q_1,\ldots,q_n=q \\ q_i\in\mathbb{Z}^+}} \binom{q}{q_1,\ldots,q_n} \times \bar{T}_{\mathbf{B}}(q_1,\ldots,q_n) \qquad (۱۵–۷)$$

که $\bar{T}_{\mathbf{B}}$ پیش از این در رابطه (۷–۲) تعریف شده است.

اثبات: برای یک جایگشت $\pi \in S_n$ می‌دانیم:

$$\big(\sum_{i=1}^{n} \ln b_{i,\pi(i)}\big)^q = \sum_{\substack{q_1,\ldots,q_n=q \\ q_i\in\mathbb{Z}^+}} \binom{q}{q_1,\ldots,q_n} \prod_{i=1}^{n} \big(\ln b_{i,\pi(i)}\big)^{q_i} \qquad (۱۶–۷)$$

اکنون با استفاده از لم ۷–۱ و تعریف (۷–۲) بدست می‌آوریم:

$$T_{\mathbf{B}}(x) = \sum_{q=0}^{\infty} \frac{x^q}{q!} \sum_{\substack{q_1,\ldots,q_n=q \\ q_i\in\mathbb{Z}^+}} \binom{q}{q_1,\ldots,q_n} \times \bar{T}_{\mathbf{B}}(q_1,\ldots,q_n) \qquad (۱۷–۷)$$

اگر در nتایی $(q_1,\ldots,q_n)$ که q_iها نامنفی‌اند و $\sum_{i=1}^{n} q_i = q$، حداقل دو صفر موجود باشد، $\big[\big(\ln b_{i,j}\big)^{q_i}\big]$ حداقل دو سطر تمام یک دارد که نتیجه می‌دهد $\bar{T}_{\mathbf{B}}(q_1,\ldots,q_n) = 0$، بنابراین nتایی‌هایی در بسط تیلور مذکور در (۷–۱۷) تاثیرگذار هستند که حداکثر شامل یک صفر باشند. از این رو در بسط تیلور $T_{\mathbf{B}}(x)$ ضرایب جملات x^q برای $q < n-1$ صفر است که نشان می‌دهد $x=0$ ریشه مکرر $T_{\mathbf{B}}$ با حداقل مرتبه $n-1$ است. ■

با استناد به قضیه ۷–۲ و لم ۷–۱ برای هر بازه محدود و هر دقت دلخواه، می‌توان $T_{\mathbf{B}}$ را توسط متناهی جمله از بسط تیلور آن تقریب زد. نتیجه می‌شود که این تقریب‌ها کلید اصلی برای محاسبه ریشه‌های $T_{\mathbf{B}}(x)$ هستند؛ سری تیلور را با n جمله تقریب می‌زنیم و سپس ریشه‌های چندجمله‌ای حاصل را بدست می‌آوریم. سپس بررسی می‌کنیم که آیا تقریب n جمله‌ای بسط تیلور به تقریب مناسبی از ریشه بدست آمده منجر می‌شود یا خیر. در صورت منفی بودن پاسخ، ریشه بدست آمده را نادیده می‌گیریم و تعداد جملات بسط تیلور را افزایش می‌دهیم تا تعداد مناسبی ریشه بدست آوریم. در قضیه زیر یک کران بالایی برای خطای تقریب‌های متناهی جمله از بسط تیلور ارائه می‌دهیم.

قضیه ۷–۳ فرض کنید $E_N(x) = \sum_{q=N+1}^{\infty} t_q x^q$ نمایشگر خطای تقریب بسط تیلور $T_{\mathbf{B}}(x)$ توسط $n+1$ جمله‌ی اول آن باشد و $M_{\mathbf{B}}$ حداکثر اندازه درایه‌های ماتریس $\ln \mathbf{B} \odot$ را نمایش دهد. برای هر x دلخواه و

$N \geq \lceil eM_{\mathbf{B}}x \rceil$ داریم:

$$|E_N(x)| \leq \frac{n^{\frac{n}{۲}} \big(eM_{\mathbf{B}}x\big)^{N+۱}}{\sqrt{۲\pi}(N+۱)^{N+۱/۵-n}\big(۱ - \frac{eM_{\mathbf{B}}x}{N+۱}\big)} \tag{۱۸–۷}$$

اثبات: با استفاده از نامساوی هادامارد خواهیم داشت:

$$\begin{aligned} \big|\bar{T}_{\mathbf{B}}(q_۱, \ldots, q_n)\big| &= \Big|\det\Big(\big[(\ln b_{i,j})^{q_i}\big]\Big)\Big| \leq \prod_{i=۱}^{n}\Big(\sum_{j=۱}^{n} |\ln b_{i,j}|^{۲q_i}\Big)^{۰/۵} \\ &\leq \prod_{i=۱}^{n} \sqrt{n} M_{\mathbf{B}}^{q_i} = n^{\frac{n}{۲}} M_{\mathbf{B}}^{\sum_{i=۱}^{n} q_i} \end{aligned} \tag{۱۹–۷}$$

دقت کنید:

$$\begin{aligned} |E_N(x)| &= \Big|\sum_{q=N+۱}^{\infty} \frac{x^q}{q!} \sum_{\substack{q_۱,\ldots,q_n=q \\ q_i \in \mathbb{Z}^+}} \binom{q}{q_۱, \ldots, q_n} \times \bar{T}_{\mathbf{B}}(q_۱, \ldots, q_n)\Big| \\ &\leq \sum_{q=N+۱}^{\infty} \frac{x^q}{q!} \sum_{\substack{q_۱,\ldots,q_n=q \\ q_i \in \mathbb{Z}^+}} \binom{q}{q_۱, \ldots, q_n} \times \Big|\bar{T}_{\mathbf{B}}(q_۱, \ldots, q_n)\Big|. \end{aligned} \tag{۲۰–۷}$$

در نتیجه با استفاده از (۷–۱۹) داریم:

$$\begin{aligned} |E_N(x)| \leq & \sum_{q=N+۱}^{\infty} \frac{x^q}{q!} \sum_{\substack{q_۱,\ldots,q_n=q \\ q_i \in \mathbb{Z}^+}} \binom{q}{q_۱, \ldots, q_n} n^{\frac{n}{۲}} M_{\mathbf{B}}^q \\ = & \; n^{\frac{n}{۲}} \sum_{q=N+۱}^{\infty} \frac{\big(M_{\mathbf{B}}x\big)^q q^n}{q!} \end{aligned} \tag{۲۱–۷}$$

به کمک $n! > \big(\frac{n}{e}\big)^n \sqrt{۲\pi n}$ بدست می‌آوریم

$$\begin{aligned} |E_N(x)| &\leq \frac{n^{\frac{n}{۲}}(eM_{\mathbf{B}}x)^{n-۰/۵}}{\sqrt{۲\pi}} \sum_{q=N+۱}^{\infty} \Big(\frac{eM_{\mathbf{B}}x}{q}\Big)^{q+۰/۵-n} \\ &\leq \frac{n^{\frac{n}{۲}}(eM_{\mathbf{B}}x)^{n-۰/۵}}{\sqrt{۲\pi}} \sum_{q=N+۱}^{\infty} \Big(\frac{eM_{\mathbf{B}}x}{N+۱}\Big)^{q+۰/۵-n} \end{aligned} \tag{۲۲–۷}$$

از آنجا که فرض کردیم $N \geq \lceil eM_{\mathbf{B}}x \rceil$ می‌دانیم $\frac{eM_{\mathbf{B}}x}{N+۱} < ۱$ و در نتیجه :

$$\sum_{q=N+۱}^{\infty} \Big(\frac{eM_{\mathbf{B}}x}{N+۱}\Big)^{q+۰/۵-n} = \Big(\frac{eM_{\mathbf{B}}x}{N+۱}\Big)^{N+۱/۵-n} \frac{۱}{۱ - \frac{eM_{\mathbf{B}}x}{N+۱}} \tag{۲۳–۷}$$

به این ترتیب اثبات کامل می‌شود. ■

شایان ذکر است که با ضرب کردن و یا تقسیم کردن درایه‌های $\mathbf{B}$ بر یک عدد ثابت، تغییری در ریشه‌های $T_{\mathbf{B}}(x)$ ایجاد نمی‌شود. با این حال، $M_{\mathbf{B}}$ و در نتیجه کران بالایی بدست آمده برای خطا تغییر می‌کند.

آخرین مطلب قابل ذکر، رتبه ماتریس $\mathbf{A}_{n\times n}$ است که ماتریس $\mathbf{B}_{n\times n}$ از روی آن ساخته شده است. همان طور که قضیه ۷–۱ نشان می‌دهد، برای تمام مقادیر m که $n > \binom{m+k-1}{m}$ (فرض کنید $m_{\max}$ بزرگترین مقدار برای چنین mهایی باشد) ماتریس $\mathbf{A}^{\odot m}$ کماکان رتبه‌ناقص است. بنابراین اگر $\mathbf{B} = \mathbf{A}^{\odot p}$ که p عددی حقیقی است، تمامی اعضای مجموعه $\{\frac{i}{p}\}_{i=0}^{m_{max}}$ ریشه‌های $T_{\mathbf{B}}(x)$ خواهند بود. به بیان بهتر، مجموعه ریشه‌های $T_{\mathbf{B}}$ شامل تصاعدی حسابی به طول $m_{max}+1$ با جمله اولیه صفر و قدر نسبت $\frac{1}{p}$ است. در حقیقت، مشاهده این تصاعد حسابی، کلید اصلی در آشکارسازی ساختار رتبه‌ناقص استفاده شده و همچنین تخمین‌های نسبتاً دقیق از $\frac{1}{p}$ است. برای روشن شدن این موضوع، به این نکته دقت کنیم که با تقریب زدن بسط تیلور توسط متناهی جمله، ریشه‌ها به طور دقیق بدست نمی‌آیند و در نتیجه مشاهده یک الگو مانند تصاعد حسابی برای تخمین قدر نسبت، گره‌گشای کار است. همچنین طول تصاعد حسابی، تخمین جالبی از رتبه ماتریس $\mathbf{A}$ (k) ارائه می‌دهد: اگر طول تصاعد l باشد داریم:

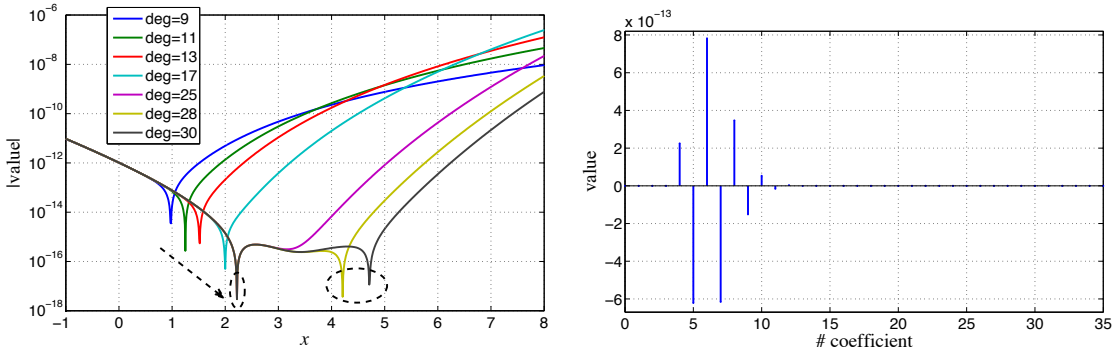
$$\binom{l+k-2}{l-1} < n \leq \binom{l+k-1}{l} \tag{۲۴–۷}$$

۵–۷ نتایج عددی

برای یک نمونه شبیه‌سازی، ماتریس زیر را به عنوان یک ماتریس رتبه‌ناقص درنظر بگیرید:

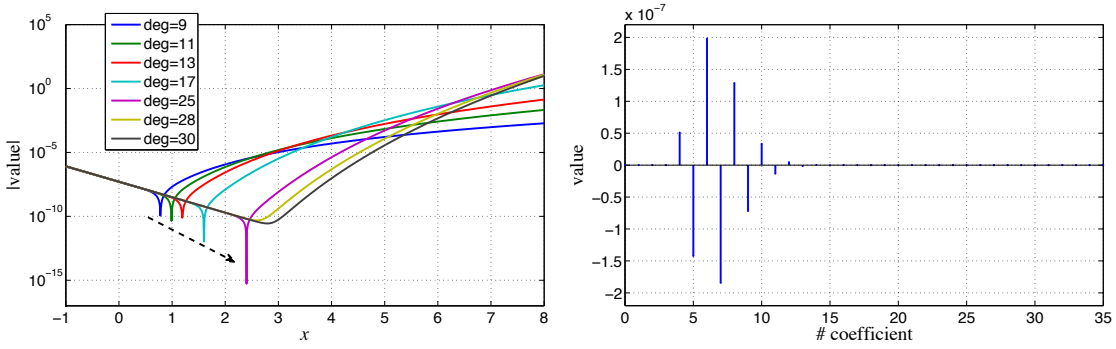
$$\mathbf{A} = \begin{bmatrix} ۲ & ۵ & ۳ & ۳ & ۵ \\ ۵ & ۱۳ & ۸ & ۷ & ۱۲ \\ ۳ & ۸ & ۵ & ۴ & ۷ \\ ۳ & ۷ & ۴ & ۵ & ۸ \\ ۵ & ۱۲ & ۸ & ۷ & ۱۳ \end{bmatrix} \tag{۲۵–۷}$$

می توان نشان داد که $rank(\mathbf{A}^{\odot i}) = i+1$ برای $i = ۱,۲,۳,۴$. مدل اعوجاج یافته $\mathbf{A}$ را در این مثال $\mathbf{B} = \mathbf{A}^{\odot \frac{۵}{۱۱}}$ فرض می‌کنیم که به معنای $p = \frac{۵}{۱۱}$ و در نتیجه $\frac{۱}{p} = \frac{۱۱}{۵} = ۲/۲$ است. برای واقعی‌تر شدن نتایج، المان‌های $\mathbf{B}$ را با نویز سفیدگوسی جمع می‌کنیم تا ماتریس $\tilde{\mathbf{B}}$ حاصل شود به نحوی که SNR آن نسبت به ماتریس $\mathbf{B}$ برابر $۱۰۰dB$ شود. شکل ۷–۲–(a) ضرایب بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ را نشان می‌دهند. همان طور که پیش‌بینی می‌شود، ضرایب فوق به سرعت میرا می‌شوند. در شکل ۷–۲–(b) مقادیر چندجمله‌ای‌های حاصل از تقریب‌های متناهی جمله بسط تیلور، در یک بازه ثابت بر حسب تعداد جمله رسم شده‌اند. از آن‌جا که بسط تیلور فوق از جمله $x^۴$ شروع می‌شود. برای سهولت در یافتن ریشه‌ها، تمامی جمله‌ها را بر $x^۴$ تقسیم کرده‌ایم. همان‌طور که نتایج نشان



(b) (a)

شکل ۷–۲: (a) ضرایب بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ و (b) مقادیر چندجمله‌ای‌های حاصل از در نظر گرفتن متناهی جمله بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ پس از حذف فاکتور $x^۴$ هنگامی که نویز جمعی در حد $\text{SNR} = ۱۰۰dB$ است.



(b) (a)

شکل ۷–۳: (a) ضرایب بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ و (b) مقادیر چندجمله‌ای‌های حاصل از در نظر گرفتن متناهی جمله بسط تیلور $T_{\tilde{\mathbf{B}}}(x)$ پس از حذف فاکتور $x^۴$ هنگامی که نویز جمعی در حد $\text{SNR} = ۵۰dB$ است.

می‌دهند، ریشه چندجمله‌ای‌های با درجه کوچکتر و یا مساوی ۱۷ از بسط اولیه (درجه ۱۳ در بسط حاصل از تقسیم بر جمله $x^۴$) با افزایش درجه به سمت مقدار مطلوب ۲٫۲ نزدیک می‌شود و هنگامی که درجه فراتر از ۲۵ رود، این ریشه با دقت مطلوبی بدست آمده و پایدار شده است (مقدار نهایی ۲٫۲۱۵۲).

به علاوه، ریشه دوم که ۴٫۴ است، برای درجه‌های بزرگتر از ۲۹ شروع به ظهور می‌کند. روند مشابهی برای حالت $\text{SNR} = ۵۰dB$ طی شده است که متعاقباً شکل‌های ۷–۳–(a) و ۷–۳–(b) حاصل شده‌اند. به طور مشابه، مشاهده می‌شود که ریشه چندجمله‌ای‌ها در درجه‌های پایین به سمت مقدار مطلوب ۲٫۲ نزدیک می‌شوند اما هنگامی که درجه از ۲۸ فراتر می‌رود، به دلیل حضور نویز، ریشه مختلط حاصل می‌شود اما کماکان در حوالی

نقطه ۲٫۲ شاهد یک مینیمم موضعی هستیم.

در حالت $\mathrm{SNR} = ۱۰۰dB$ با استفاده از تخمین $\frac{۱}{p}$ به صورت ۲٫۲۱۵۲ از ماتریس $\tilde{\mathbf{B}}$ به تخمینی از ماتریس $\mathbf{A}$ می‌رسیم که مقادیر تکین آن عبارتند از: $\{۳٫۵۹۸, ۰٫۲۰۱, ۱٫۳۹e-۵, ۲٫۹۳e-۷, ۱٫۷۹e-۹\}$. واضح است که سه مقدار تکین کوچکتر، در واقع صفر بوده‌اند که به دلیل وجود نویز و غیرایده‌آل بودن تخمین $\frac{۱}{p}$ به این صورت ظاهر شده‌اند. مقادیر تکین ماتریس $\mathbf{A}$ عبارت است از $\{۳٫۶, ۰٫۲, ۰, ۰, ۰\}$.

فصل ۸

جمع بندی و نتیجه‌گیری

در این پایان‌نامه به طور خاص به بررسی روش‌های غیرتصادفی نمونه‌برداری از سیگنال‌های تنک پرداختیم. در روش‌های خطی این عمل به کمک ماتریس حسگر صورت می‌گیرد. در این پایان‌نامه با چند روش مختلف، ساخت این ماتریس‌ها مورد بررسی قرار گرفت. به کمک کدهای متعامد نوری روش ساختی برای ماتریس‌های حسگر دودویی ارائه شد که از نظر ابعاد در حد ماتریس‌های معرفی شده توسط Devore است با این تفاوت که تنوع روش‌های ساخت بیشتر است. نامنفی‌بودن درایه‌های ماتریس، عامل بسیار محدودکننده‌ای در طراحی است؛ برای غلبه بر این مشکل به سراغ ماتریس‌های دو قطبی و مختلط رفتیم. به کمک کدهای BCH دودویی، ماتریس‌های دو قطبی $(l > j)$ ۲^{-j} $(۲^l - ۱) \times ۲^{\mathcal{O}\left(۲^{(l-j)\frac{\ln j}{j}}\right)}$ را معرفی کردیم که ضریب همدوسی کمتر از دارند و در نتیجه شرط RIP از مرتبه $k \leq ۲^j + ۱$ را ارضا می‌کنند. با کمک تعمیم‌های غیردودویی کدهای BCH این طرح به ماتریس‌های مختلط $(p^l - ۱) \times p^{\mathcal{O}\left(p^{(l-j)\frac{\ln j}{j}}\right)}$ با p اول تعمیم داده شد به طوری که درایه‌های ماتریس هم‌اندازه‌اند. در اکثر این ساختارها عملکرد ماتریس حسگر در حد ماتریس‌های تصادفی و گاهی بهتر مشاهده شده است؛ علاوه بر این، ساختار دوری ستون‌های این ماتریس‌ها که ناشی از گردشی بودن کدهای BCH است به طور موثری پیچیدگی محاسباتی در بازسازی بردار تنک از روی نمونه‌ها را کاهش می‌دهد. در ادامه، روش‌هایی برای ادغام این ماتریس‌ها و ایجاد تنوع بیشتر در ابعاد، معرفی کردیم؛ به کمک ادغام ماتریس‌ها توانستیم ماتریس‌هایی با ابعاد $p^۲ \times p^{r+۱} ۲^{\mathcal{O}\left(r \frac{\ln(\log_۲ p - \log_۲ r)}{\log_۲ p - \log_۲ r}\right)}$ (p توانی از یک عدد اول) بسازیم که تا کنون بهترین کران در ساختارهای غیرتصادفی به شمار می‌رود.

نمونه‌برداری غیرخطی برای اولین بار در این پایان‌نامه مورد بحث قرار گرفت. نشان دادیم هنگامی که

دسته سیگنال‌ها را به بردارهای تک تنک محدود کنیم، روش‌های غیرخطی وجود دارند که مستقل از حوزه تنک‌بودن سیگنال، علاوه بر فشرده‌سازی، اطلاعات موجود در سیگنال را حفظ می‌کنند. این در حالی است که روش‌های خطی قادر به چنین کاری نیستند. همچنین نشان دادیم که با استفاده از طرح‌های غیرخطی می‌توان به نرخ فشرده‌سازی بالاتر و در عین حال پیچیدگی محاسباتی کمتر در گیرنده دست یافت. البته همان‌طور که انتظار می‌رود این روش‌ها نسبت به نویز ناپایدارند و تنها در کاربردهایی که نویز جمعی وجود ندارد (مانند شبکه‌های داده و نوری) می‌توانند جایگزین مناسبی برای روش‌های خطی باشند.

از مفاهیم عمیق در مبحث نمونه‌برداری فشرده، تعاریف تنک‌بودن و فشرده‌پذیری است؛ تمام نتایج این مبحث بر پایه تنک‌بودن سیگنال‌های ورودی استوار است. با وجود آن که این مفاهیم در حالت گسسته و بعد متناهی تعاریف قابل قبولی دارند، در حالت بعد نامتناهی وضعیت کاملاً متفاوت است. در این پایان‌نامه نشان دادیم که این مفاهیم را چه برای سیگنال‌های یقینی و چه تصادفی، می‌توان به بعد نامتناهی تعمیم داد. تعاریف ارائه شده، بر مبنای رفتار حدی دنباله‌های منقطع شده استوار است و نتایج بعد متناهی را در بر می‌گیرد. همچنین به کمک دو قضیه، روشی برای متمایز کردن توزیع احتمال‌هایی که دنباله تصادفی و i.i.d فشرده‌پذیر تولید می‌کنند معرفی گردید.

در انتهای پایان‌نامه نیز یک مسأله تنک دو بعدی را بررسی کردیم. ماتریس‌ها، تعمیم‌هایی از بردارها به شمار می‌روند، حال آن که در تعمیم مباحث نمونه‌برداری فشرده به ماتریس‌ها، بر خلاف بردارها که تنک‌بودن براساس درایه‌ها تعریف می‌شود، تنک‌بودن براساس رتبه ماتریس بیان می‌شود. در این‌جا نشان دادیم که چگونه می‌توان یک ماتریس کم‌رتبه را که درایه‌هایش به طور یکسانی تحت تاثیر یک اعوجاج غیرخطی قرار گرفته‌اند (و در نتیجه ماتریس دیگر کم رتبه نیست) بازیابی کرد. از نتایج جالب توجه آن است که اگر رتبه ماتریس به اندازه کافی کوچک باشد ودرایه‌ها به توان یک عدد صحیح و مثبت رسیده باشند، ماتریس حاصل کماکان رتبه‌ناقص است اما خواص متمایز کننده‌ای نسبت به یک ماتریس رتبه‌ناقص حالت کلی دارد. این حالت، تعمیمی از مساله کاهش پهنای باند در سیگنال‌های یک بعدی است که رابطه تنگاتنگی با مبحث نمونه‌برداری فشرده دارد. در حالتی که توان عدد صحیح نباشد، ماتریس نهایی به احتمال زیاد رتبه کامل است اما می‌توان با بررسی خواص آن وجود، ماتریس کم رتبه را تایید و یا تکذیب کرد.

پژوهش‌هایی که از دید نگارنده در ادامه کارهای این پایان‌نامه قرار دارند عبارتند از:

۱. ارائه ماتریس‌های یقینی با ابعاد نزدیک‌تر به ماتریس‌های تصادفی. استفاده از گراف‌های مربوط به کدهای LDPC روشی امیدوار کننده به نظر می‌رسد (نحوه استفاده از گراف‌ها در ساخت ماتریس به طور مختصر در فصل ۱ شرح داده شد).

۲. بررسی زیرماتریس‌های (انتخاب چند سطر) ماتریس DFT به منظور یافتن بهترین زیرماتریس از نظر برآورده کردن شرط RIP. با کمک تحلیل‌های تصادفی، وجود زیرماتریس‌هایی که شرط RIP را با ضریب مناسب ارضا می‌کنند به اثبات رسیده است، اما ساختار غیرتصادفی مناسبی هنوز ارائه نشده است.

۳. در شبیه‌سازی‌های ارائه شده در این پایان‌نامه، تنها روش بازسازی OMP مورد بررسی قرار گرفته است. استفاده از سایر روش‌ها مانند BP، CoSaMP، LASSO و ... از جمله کارهای آینده‌اند.

۴. تحلیل حساسیت روش‌های نمونه‌برداری غیرخطی غیرتصادفی نسبت به نویز جمعی و ایجاد بده-بستان بین تعداد نمونه و حساسیت به نویز.

۵. معرفی و بررسی روش‌های نمونه‌برداری غیرخطی تصادفی (در این پایان‌نامه روشی غیرتصادفی ارائه شد).

۶. تعمیم قضایای نمونه‌برداری فشرده مانند مرتبه تعداد نمونه لازم و شرط کافی بازسازی (RIP) به سیگنال‌های بینهایت بعدی و پیوسته به کمک تعاریف تعمیم‌یافته در این پایان‌نامه.

۷. کاهش حساسیت به نویز در روش‌های غیرخطی بازیابی ماتریس کم رتبه. پیش از بازیابی ممکن است بتوان با اعمال یک تابع مناسب بر روی درایه‌ها (مانند حالتی که در نمونه‌برداری غیرخطی مطالعه شد)، حساسیت را کاهش داد.

۸. بازیابی ماتریس کم‌رتبه اعوجاج یافته هنگامی که تعدادی از درایه‌های آن موجود نباشند (در مباحث فصل ۵ تمامی درایه‌ها معلوم فرض شده‌اند).

۹. بررسی کاربردهای ماتریس‌های یقینی با ساختار Devore و BCH در ارتباطات نوری هم‌زمان و بی‌سیم غیرهم‌زمان.

پیوست الف

محاسبه $\tilde{k}$

در قضیه ۴-۲ نشان دادیم که $\tilde{k}$ برابر است با تعداد دنباله‌های باینری به طول $\tilde{m}$ که میان هر دو ۱ حداقل $\tilde{m} - l - ۱$ صفر به صورت چرخشی وجود داشته باشد. برای محاسبه این عدد، فرض کنید $\tau_b^{(a)}$ نمایانگر تعداد دنباله‌های باینری به طول b باشد که در آن‌ها میان هر دو ۱ متوالی حداقل aتا صفر قرار داشته باشد. همچنین فرض کنید $\kappa_b^{(a)}$ تعداد دنباله‌های باینری به طول b را نشان دهد که بین هر دو ۱ متوالی آن‌ها حداقل aتا صفر (نه لزوماً به صورت چرخشی) یافت شود. در ادامه، ابتدا به محاسبه $\kappa_b^{(a)}$ می پردازیم و سپس ارتباط میان $\tau_b^{(a)}$ و $\kappa_b^{(a)}$ را بدست می‌آوریم.

دو نوع متفاوت دنباله باینری در $\kappa_b^{(a)}$ شمرده می‌شوند:

۱. دنباله‌هایی که به صفر ختم می‌شوند؛ با حذف بیت انتهایی این دنباله‌ها، به دنباله‌ای با طول $b - ۱$ با خاصیت مشابه می‌رسیم. همچنین هر دنباله باینری به طول $b - ۱$ با خاصیت مطلوب را می‌توان با افزودن یک بیت صفر به انتهای آن، به دنباله‌ای به طول b با خاصیت مورد نظر تبدیل کرد. در نتیجه تعداد این دنباله‌ها $\kappa_{b-۱}^{(a)}$ است.

۲. دنباله‌هایی که به یک ختم می‌شوند؛ این بدان معناست که $a + ۱$ بیت انتهایی دنباله به صورت $۱,\underbrace{۰,\ldots,۰}_{a}$ است. مشابه حالت ۱، هر دنباله باینری به طول $b - a - ۱$ با خاصیت مطلوب را می‌توان با افزودن بلوک $۱,\underbrace{۰,\ldots,۰}_{a}$ به انتهای آن، به دنباله‌ای به طول b با همین خاصیت تبدیل کرد. در نتیجه، تعداد این دنباله‌های باینری $\kappa_{b-a-۱}^{(a)}$ است.

با درنظرگرفتن دو حالت قبل، دنباله بازگشتی زیر بدست می‌آید:

$$\kappa_b^{(a)} = \kappa_{b-۱}^{(a)} + \kappa_{b-a-۱}^{(a)} \qquad \text{(۱–الف)}$$

از آنجا که برای $b \le a+۱$ حداکثر یک بیت برابر با ۱ در دنباله وجود دارد، داریم:

$$۱ \le b \le a+۱ : \ \kappa_b^{(a)} = b+۱ \qquad \text{(۲–الف)}$$

با استفاده از (۱–الف)، به جای آخرین شرط اولیه ($\kappa_{a+۱}^{(a)} = a+۲$)، می‌توانیم شرط $\kappa_۰^{(a)} = ۱$ را اضافه کنیم. اکنون تبدیل $\mathcal{Z}$ یک طرف دنباله $\kappa_b^{(a)}$ که به صورت

$$\kappa^{(a)}(z) = \sum_{b=۰}^{\infty} \kappa_b^{(a)} z^{-b}, \qquad \text{(۳–الف)}$$

تعریف می‌شود، عبارت است از:

$$\kappa^{(a)}(z) = \frac{۱}{۱-z^{-۱}} \cdot \frac{۱-z^{-(a+۱)}}{۱-z^{-۱}-z^{-(a+۱)}} \qquad \text{(۴–الف)}$$

بنابراین، نرخ رشد $\kappa_b^{(a)}$ بر حسب b از مرتبه γ^b است که γ بزرگترین (از نظر قدر مطلق) ریشه $f(z) = z^{a+۱} - z^a - ۱$ است. ازآنجا که $f(۱) \cdot f(۲) < ۰$ این چندجمله‌ای دربازه (۲ , ۱) دست‌کم یک ریشه دارد؛ این ریشه را با γ نشان می‌دهیم. در حقیقت این ریشه بزرگترین ریشه $f(z)$ است که در این‌جا به اثبات آن نمی‌پردازیم، اما در صورتی که $f(z)$ ریشه بزرگتری داشته باشد، نتایجی که در ادامه بدست می‌آید می‌تواند به عنوان کران پایینی برای نرخ‌های بدست‌آمده به‌شمار رود. اکنون در پی تقریب مناسبی از γ هستیم که:

$$۱ < \gamma < ۲ \ , \quad f(\gamma) = \gamma^{a+۱} - \gamma^a - ۱ = ۰ \qquad \text{(۵–الف)}$$

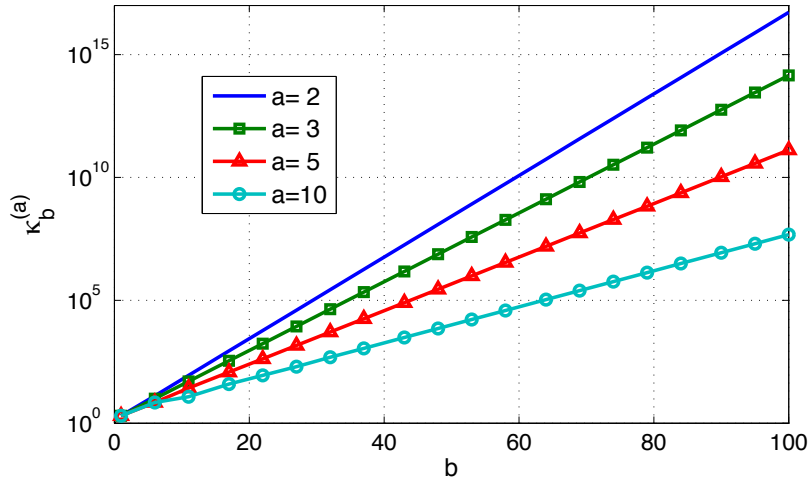
از آن‌جا که $۱ < \gamma < ۲$ می‌توانیم فرض کنیم $\gamma = ۱ + \frac{۱}{\delta}$ که $\delta > ۱$:

$$\gamma^{a+۱} - \gamma^a = ۱ \ \Rightarrow \ \left(۱ + \frac{۱}{\delta}\right)^a = \delta \qquad \text{(۶–الف)}$$

در پیوست ب نشان می‌دهیم که $\delta > a^{۰/۸۷}$ و در نتیجه:

$$\gamma^{\frac{a}{\ln a}} = \left(۱ + \frac{۱}{\delta}\right)^{\frac{a}{\ln a}} = \delta^{\frac{۱}{\ln a}} > a^{\frac{۰/۸۷}{\ln a}} = e^{۰/۸۷} > ۲ \qquad \text{(۷–الف)}$$

حال ارتباط بین $\tau_b^{(a)}$ و $\kappa_b^{(a)}$ را بررسی می‌کنیم. بنا به تعریف، تمام دنباله‌های شمرده شده در $\tau_b^{(a)}$، در $\kappa_b^{(a)}$ نیز شمرده می‌شوند و در نتیجه $\tau_b^{(a)} \le \kappa_b^{(a)}$ همچنین اگر به انتهای یک دنباله شمرده‌شده در $\kappa_{b-a}^{(a)}$، a صفر اضافه



شکل ۱–الف: مقادیر بدون تقریب $\kappa_b^{(a)}$ در چند انتخاب متفاوت از a و b.

کنیم، دنباله حاصل‌شده در شمارش $\tau_b^{(a)}$ محسوب می‌شود. پس $\kappa_{b-a}^{(a)} \leq \tau_b^{(a)}$. با ترکیب این دو نامساوی داریم:

$$\mathcal{O}(\gamma^{b-a}) \leq \tau_b^{(a)} \leq \mathcal{O}(\gamma^b) \qquad \text{(۸–الف)}$$

با ادغام رابطه فوق و نتیجه ذکر شده در (۷–الف) بدست می‌آوریم:

$$\left(\tau_b^{(a)}\right)^{\frac{a}{\ln a}} \gtrapprox \mathcal{O}\left(۲^{b-a}\right) \qquad \text{(۹–الف)}$$

تعبیر رابطه فوق در مورد $\tilde{k}$ عبارت است از:

$$\tilde{k} = \tau_{\tilde{m}}^{(\tilde{m}-l-۱)} \gtrapprox \mathcal{O}\left(۲^{(l+۱)\frac{\ln \tilde{m}-l-۱}{\tilde{m}-l-۱}}\right) \qquad \text{(۱۰–الف)}$$

دقت کنید که $\tilde{m} - l - ۱ \leq \log_۲ k$ که k حداکثر مرتبه RIP است که توسط استدلالات مبتنی بر ضریب همدوسی تضمین می‌شود. به همین دلیل برای ماتریس‌های دو قطبی ساخته‌شده خواهیم داشت:

$$m = ۲^{\tilde{m}} - ۱ < ۲^{\tilde{m}-l-۱}۲^{l+۱} \lessapprox \mathcal{O}\left(k\left(\log_۲ n\right)^{\frac{\log_۲ k}{\ln \log_۲ k}}\right) \qquad \text{(۱۱–الف)}$$

شکل ۱–الف رفتار حدی $\kappa_b^{(a)}$ را برای مقادیر مختلف a و b نشان می‌دهد. به طور مثال، شکل ۱–الف نشان می‌دهد که $\kappa_b^{(۵)} \approx ۱/۶۶ \times ۱/۲۸۵^b$ در حالی که تقریب‌های استفاده‌شده حاکی از $\mathcal{O}(۱/۲۵^b) \geq \kappa_b^{(۵)}$ است.

پیوست ب

محاسبه δ

در این پیوست نشان می‌دهیم که اگر $\delta = (1 + \frac{1}{\delta})^a$ که $a \in \mathbb{N}$، در این صورت $\delta > a^{0/7}$. برای این منظور لم زیر را بیان می کنیم:

لم ۱–ب تابع $f(x) = x^{0/3} - 0/5x^{-0/4} - 0/7 \ln x$ برای $x > 0$ همواره مثبت است.

اثبات: واضح است که $\lim_{x\to+\infty} f(x) = +\infty$. نشان می‌دهیم $f'(x)$ (مشتق f) تنها یک ریشه در $[0, \infty)$ دارد که مینیمم موضعی f را تولید می‌کند:

$$f'(x) = 0/3x^{-0/7} + 0/2x^{-1/4} - 0/7x^{-1} = frac3(x^{0/1})^7 - 7(x^{0/1})^4 + 210x^{1/4} \qquad (۱–ب)$$

تنها ریشه مثبت چندجمله‌ای $3y^7 - 7y^4 + 2$ برابر با $y \approx 1/277$ است؛ در نتیجه حداقل مقدار f روی محور مثبت در نقطه $x \approx 1/277^{10} \approx 11/532$ حاصل می‌شود. با محاسبه مقدار تابع در این نقطه داریم $f(11/532) \approx 0/18 > 0$. ■

حال با استفاده از لم فوق برای $x \geq 1$ داریم:

$$0 < f(x) = x\left(x^{-0/7} - 0/5x^{-1/4}\right) - \ln x^{0/7} \leq x \ln\left(1 + x^{-0/7}\right) - \ln x^{0/7} \qquad (۲–ب)$$

بنابراین:

$$e^{\ln x^{0/7}} < e^{x \ln\left(1+x^{-0/7}\right)} \quad \Rightarrow \quad \frac{\left(x^{0/7}\right)^{x+1}}{\left(1 + x^{0/7}\right)^x} < 1 \qquad (۳–ب)$$

دقت کنید که تابع $\psi(x) = \frac{x^{a+1}}{(1+x)^a} = x(1 - \frac{1}{1+x})^a$ روی محور مثبت، اکیداً صعودی است (هر دوی x و $1 - \frac{1}{1+x}$ توابعی صعودی هستند). با استفاده از رابطه $\delta = (1 + \frac{1}{\delta})^a$ می‌دانیم که $\psi(\delta) = 1$؛ از سوی دیگر، با قراردادن $x = a$ در (۳–ب) بدست می‌آوریم $\psi(a^{0.1}) < 1$ که نشان می‌دهد $\delta > a^{0.1}$.

مراجع

[1] D. Achlioptas, "Database-friendly random projections," in *ACM SIGACT-SIGMOD-SIGART Symp. on Principles of Database Systems*, 2001, pp. 274–281.

[2] A. Amini, A. Karbasi, F. Marvasti, M. Vetterli, and M. Unser, "Low-rank matrix approaximation using point-wise operators," *Submitted to IEEE Trans. Inform. Theory.*

[3] A. Amini and F. Marvasti, "Deterministic construction of binary, bipolar and ternary compressed sensing matrices," *To appear in IEEE Trans. Inform. Theory.*

[4] A. Amini and F. Marvasti, "Convergence analysis of an iterative method for the reconstruction of multi-band signals from their uniform and periodic nonuniform samples," *Sampling Theory in Signal and Image Processing*, vol. 7, no. 2, pp. 113–130, Jan. 2008.

[5] A. Amini and F. Marvasti, "Limits of deterministic compressed sensing considering arbitrary orthonormal basis for sparsity," in *Sampling Theory and Applications (SAMPTA2009)*, May 2009.

[6] A. Amini, V. Montazerhodjat, and F. Marvasti, "Deterministic rip-fulfilling matrices using p-ary block codes," *Conditionally accepted in IEEE Trans. Sig. Proc.*

[7] A. Amini, V. Montazerhodjat, and F. Marvasti, "RIP-fulfilling complex-valued matrices," in *Intern. Conf. on Communications (ICC2010)*, May 2010.

[8] A. Amini, M. Unser, and F. Marvasti, "Compressibility of deterministic and random infinite sequences," *Submitted to IEEE Trans. Sig. Proc.*

[9] L. Applebaum, S. D. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Applied and Computational Harmonic Analysis*, vol. 26, no. 2, pp. 283–290, March 2009.

[10] R. Baraniuk, "A lecture on compressive sensing," *IEEE Sig. Proc. Magazine*, vol. 241, pp. 118–121, July 2008.

[11] R. Baraniuk, M. Davenport, R. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constr. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[12] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Scie.*, vol. 2, no. 1, pp. 183–202, March 2009.

[13] M. W. Berry, Z. Drmac, and E. R. Jessup, "Matrices, vector spaces, and information retrieval," *SIAM Rev.*, vol. 41, no. 2, pp. 335–362, 1999.

[14] T. Blumensath and M. E. Davies, "Compressed sensing and source separation," in *Conf. Independent Component Analysis and Signal Separation*, 2007.

[15] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *Fourier Anal. and App.*, vol. 14, no. 5-6, pp. 629–654, April 2007.

[16] J. Bobin, J. L. Starck, and R. Ottensamer, "Comressed sensing in astronomy," *IEEE Journal of Selected Topics in Sig. Proc.*, vol. 2, no. 5, pp. 718–726, Oct. 2008.

[17] J. L. Brown, "Sampling extensions for multiband signals," *IEEE Trans. Acoust. Speech, Signal Proc.*, vol. 33.

[18] J. L. Brown, "Sampling rate reduction in multichannel processing of bandpass signals," *J. Acoust. Soc. Amer*, vol. 71, no. 2, pp. 378–383, 1982.

[19] R. Calderbank, S. Howard, and S. Jafarpour, "Construction of a large class of deterministic sensing matrices that satisfy a statistical isometry property," *IEEE Journal of Selected Topics in Sig. Proc.*, vol. 4, no. 2, pp. 358–374, April 2010.

[20] R. Calderbank, S. Jafarpour, and R. Schapire, "Compressed learning: Universal sparse dimensionality reduction and learning in the measurement domain," Tech. Rep., 2009.

[21] E. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.

[22] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Prob.*, vol. 23, no. 3, pp. 969–985, Oct. 2007.

[23] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[24] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[25] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[26] E. J. Candès, "Compressive sampling," in *Proc. International Congress of Mathematicians*, 2006, vol. 3, pp. 1433–1452.

[27] V. Cevher, "Learning with compressible priors," in *Proc. Neural Information Processing Systems (NIPS)*, 2008, vol. Vancouver, B.C., Canada.

[28] V. Cevher, "Approximate distributions for compressible signals," in *Proc. IEEE Information Theory Workshop*, Oct. 2009.

[29] R. H. Chan, Y. Dong, and M. Hintermuller, "An efficient two-phase L(1)-TV method for restoring blurred images with impulse noise," *IEEE Trans. Image Proc.*, vol. 19, no. 7, pp. 1731–1739, July 2010.

[30] J. Chen, K. Yao, and R. E. Hudson, "Source localization and beamforming," *IEEE Signal Proc. Magazine*, vol. 19, no. 2, pp. 30–39, May 2002.

[31] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, pp. 129–159, 2001.

[32] K. L. Clarkson and Peter W. Shor, "Applications of random sampling in computational geometry," *Springer, Discrete and Computational Geometry*, vol. 4, no. 1, pp. 387–421, 1989.

[33] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k-term approximation," *J. Amer. Math. Soc.*, vol. 22, pp. 211–231, 2009.

[34] A. J. Coulson, "A generalization of nonuniform bandpass sampling," *IEEE Trans. Sig. Proc.*, vol. 43, no. 3, pp. 694–704, 1995.

[35] I. Daubechies, R. DeVore, M. Fornasier, and S. Gunturk, "Iteratively re-weighted least squares minimization for sparse recovery," *Comm. on Pure and Applied Math.*, vol. 63, no. 1, pp. 1–38, Jan. 2010.

[36] M. R. de Prony, "Essai expèrimentalle et analytique," *J. Ècole Polytech. Paris*, vol. 1, pp. 24–76, 1795.

[37] E. Van den berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," Tech. Rep., Department of Computer Science, University of British Columbia, 2008.

[38] R. A. DeVore, "Deterministic construction of compressed sensing matrices," *Journal of Complexity*, vol. 23, no. doi:10.1016/j.jco.2007.04.002, pp. 918 –925, March 2007.

[39] C. Ding and C. Xing, "Several classes of ($۲^{m-۱}$, w, ۲) optical orthogonal codes," *Discr. Applied Math.*, vol. 128, no. 1, pp. 103 –120, May 2003.

[40] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[41] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal $\ell_۱$-norm solution is also the sparsest solution," *Tech. Rep.*, vol. 20, no. 1, pp. 33–61, 2004.

[42] P. Drineas, A. Javed, M. Magdon-Ismail, G. Pandurangant, R. Virrankoski, and A. Savvides, "Distance matrix reconstruction from incomplete distance information for sensor network localization," in *Sensor and Ad-Hoc Communications and Networks Conference (SECON)*, Sept. 2006, vol. 2, pp. 536–544.

[43] M. F. Duarte and R. G. Baraniuk, "Kronecker compressive sensing," *http://www.math.princeton.edu/ mduarte/images/KCS-TIP09.pdf*, 2010.

[44] J. H. Ender, "On compressive sensing applied to radar," *Signal Processing*, vol. 90, no. 5, pp. 1402–1414, May 2010.

[45] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, John Wiley, 2 edition, 1991.

[46] M. Figueiredo, R. Nowak, and S. Wright, "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," *IEEE Trans. Selected Topics in Sig. Proc.*, vol. 1, no. 4, pp. 586–597, Dec.

[47] A. K. Fletcher, S. Rangan, and V. K. Goyal, "On the rate-distortion performance of compressed sensing," in *IEEE Int. Conf. Acoustic, Speech and Sig. Proc. (ICASSP2007)*, Apr. 2007, vol. 3, pp. 885–888.

[48] S. Foucart and M. J. Lai, "Sparsest solutions of underdetermined linear systems via ℓ_q minimization for $\circ < q \leq ١$," *App. and Comp. Harmonic Analysis*, vol. 26, no. 3, pp. 395–407, May 2009.

[49] W. Gautschi and G. Inglese, "Lower bounds for the condition number of vandermonde matrices," *Springer Numerische Mathematik*, vol. 52, pp. 241–250, 1988.

[50] A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin, "Algorithmic linear dimension reduction in the l1-norm for sparse vectors," in *Allerton Conf. on Comm.*, 2006.

[51] A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin, "One sketch for all: Fast algorithms for compressed sensing," in *ACM STOC2007*, Nov. 2007, vol. 15, pp. 237–246.

[52] V. K. Goyal, A. K. Fletcher, and S. Rangan, "Compressive sampling and lossy compression," *IEEE Sig. Proc. Magazine*, vol. 25, no. 2, pp. 48–56, March 2008.

[53] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[54] S. D. Howard, A. R. Calderbank, and S. J. Searle, "A fast reconstruction algorithm for deterministic compressive sensing using second order reed-muller codes," in *IEEE Conf. on Inform. Sciences and Systems (CISS2008)*, 2008.

[55] P. Indyk, "Explicit constructions for compressed sensing of sparse signals," in *ACM-SIAM symp. on Discrete Algorithms*, 2008, pp. 30–33.

[56] P. Indyk and R. Motwani, "Approximate nearest neighbours: towards removing the curse of dimensionality," in *Symp. on Theory of Computing*, 2001, pp. 604–613.

[57] S. Jafarpour, W. Xu, B. Hassibi, and R. Calderbank, "Efficient and robust compressed sensing using optimized expander graphs," *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 4299–4308, Sept. 2009.

[58] S. M. Johnson, "A new upper bound for error-correcting codes," *IRE Trans. Inform. Theory*, vol. 8, pp. 203–207, 1962.

[59] W. B. Johnson and J. Lindenstrauss, "Extensions of lipschitz mappings into a hilbert space," in *Conf. in Modern Analysis and Probability*, 1984, pp. 189–206.

[60] S. Jokar and V. Mehrmann, "Sparse solutions to under-determined Kronecker product systems," *Linear Algebra and its Applications*, vol. 431, no. 12, pp. 2437–2447, Dec. 2009.

[61] I.T. Jolliffe, *Principal Component Analysis*, Springer, New York, 1986.

[62] A. Karbasi, S. Oh, R. Parhizkar, and M. Vetterli, "Ultrasound tomography calibration using structured matrix completion," in *International Congress on Acoustics (ICA2010)*, 2010.

[63] B. S. Kashin, "Diameters of some finite-dimensional sets and classes of smooth functions," *Izv. Akad. Nauk SSSR Ser. Mat.*, vol. 41, pp. 334–351, 1977.

[64] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from noisy entries," *arXiv:0906.2027v1*, 2009.

[65] R. Lepage, M. Woodroofe, and J. Zinn, "Convergence to a stable distribution via order statistic," *Annals of Probability*, vol. 9, no. 4, pp. 624–632, 1981.

[66] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*, Prentice Hall: Englewood Cliffs, 2 edition, 2004.

[67] G. Lorentz, M. Golitschek, and Y. Makovoz, *Constructive Approximation: Advanced Problems*, vol. 304, Springer, Berlin, 1996.

[68] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Panly, "Compressed sensing mri," *IEEE Sig. Proc. Magazine*, vol. 25, no. 2, pp. 72–82, March 2008.

[69] F. Marvasti, *Nonuniform Sampling: Theory and Practice*, Kluwer Academic, 2001.

[70] F. Marvasti, A. Amini, F. Haddadi, M. Soltanolkotabi, A. Aldroubi, S. Sanei, and J. Chambers, "A unified approach to sparse signal processing," *arXiv:0902.1853v1*, 2009.

[71] F. Marvasti, M. Hasan, M. Eckhart, and S. Talebi, "Efficient algorithms for burst error recovery using fft and other transform kernels," *IEEE Trans. on Sig. Proc.*, vol. 47, no. 4, pp. 1065–1075, 1999.

[72] F. Marvasti and A. Jain, "Zero-crossings, bandwidth compression and restoration of bandlimited signals distorted by nonlinear systems," *Journal of the Optical Society of America*, vol. 3, no. 5, pp. 651–654, 1986.

[73] H. Mohimani, M. Babaie-zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed $\ell_\circ$ norm," *IEEE Trans. Sig. Proc.*, vol. 57, no. 1, pp. 289–301, Jan. 2009.

[74] D. Needell and J. A. Tropp, "Cosamp: iterative signal recovery from incomplete and inaccurate samples," *App. and Comp. Harmonic Analysis*, vol. 26, pp. 301–321, May 2009.

[75] H. Nyquist, "Certain topics in telegraph transmission theory," *AIEE Trans.*, vol. 47, pp. 617–644, 1928.

[76] S. Oh, A. Karbasi, and A. Montanari, "Sensor network localization from local connectivity : Performance analysis for the mds-map algorithm," in *IEEE Information Theory Workshop (ITW2010)*, 2010.

[77] T. Park and G. Casella, "The Bayesian lasso," *Journal of the American Statistical Association*, vol. 103, pp. 681–686, June 2008.

[78] V.F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophys. J. Roy. Astron. Soc.*, vol. 33, pp. 347–366, 1973.

[79] T. Rappaport, *Wireless Communications: Principles and Practice*, Prentice Hall, 2001.

[80] Y. Rivenson and A. Stern, "Compressed imaging with separable sensing operator," *IEEE Sig. Proc. Letters*, vol. 16, no. 6, pp. 449–452, June 2009.

[81] J. A. Salehi, "Code division multiple-access techniques in optical fiber networks- part I: Fundamental principles," *IEEE Tran. Comm.*, vol. 37, no. 8, pp. 824–833, Aug. 1989.

[82] G. Samorodnitsky and M. S. Taqqu, *Stable non-Gaussian Random Processes*, Chapman & Hall/CRC, 1994.

[83] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagation*, vol. 34, pp. 276–280, March 1986.

[84] C. E. Shannon, "Communication in the presence of noise," *IRE*, vol. 37, pp. 10–21, 1949.

[85] M. Soltanalian, M. Soltanolkotabi, A. Amini, and F. Marvasti, "A practical sparse channel estimation for current ofdm standards," in *International Conf. on Telecommunications (ICT2009)*, 2009.

[86] M. Soltanolkotabi, A. Amini, and F. Marvasti, "OFDM channel estimation based on adaptive thresholding for sparse signal detection," in *European Sig. Proc. Conf. (EUSIPCO2009)*, 2009.

[87] T. Strohmer and R. W. Heath, "Grassmannian frames with applications to coding and communication," *Applied and Computational Harmonic Analysis*, vol. 14, no. 3, pp. 257–275, May 2003.

[88] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.

[89] J. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. on Inform. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[90] J. Tropp, "Recovery of short linear combinations via ℓ_1 minimization," *IEEE Trans. on Inform. Theory*, vol. 90, no. 4, pp. 1568–1570, July 2005.

[91] J. Tropp and A.C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[92] Matthew Turk and Alex Pentland, "Eigenfaces for recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[93] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Sig. Proc.*, vol. 50, no. 6, pp. 1417–1428, June 2002.

[94] C. Vonesch and M. Unser, "Fast iterative thresholding algorithm for wavelet-regularized deconvolution," in *Conf. Mathematical Methods: Wavelet XII*, 2007, vol. 6701.

[95] D. P. Wipf and B. D. Rao, "Sparse Baysian learning for basis selection," *IEEE Trans. on Sig. Proc.*, vol. 52, no. 8, pp. 2153–2164, Aug. 2004.

[96] D. L. Donoho Y. Tsaig, "Extensions of compressed sensing," Tech. Rep., Department of Statistics, Stanford University, 2004.

[97] S. Zahedpour, S. Feizi-Khankandi, A. Amini, M. Ferdosizadeh, and F. Marvasti, "Impulsive noise cancellation based on soft decision and recursion," *IEEE Trans. on Instrumentation and Measurement*, vol. 58, no. 8, pp. 2780–2790, Aug. 2009.

[98] A. I. Zayed, *Advances in Shannon's Sampling Theory*, CRC press, 1993.

[99] Y. Zhang, J. Schneider, and A. Dubrawski, "Learning compressible models," in *SIAM International Conference on Data Mining*, April 2010.

ABSTRACT

The emerging field of compressed sensing deals with the techniques of combining the two blocks of sampling and compression into a single unit without compromising the performance. Clearly, this is not feasible for any general signal; however, if we restrict the signal to be sparse, it becomes possible.

There are two main challenges in compressed sensing, namely the sampling process and the reconstruction methods. In this thesis, we will focus only on the deterministic sampling process as opposed to the random sampling. The sampling methods discussed in the literature are mainly linear, i.e., a matrix is used as the sampling operator. Here, we first consider linear sampling methods and introduce some deterministic designs. The constructed matrices are derived from OOC, BCH and non-binary BCH codes. The cyclic property of BCH codes enables us to implement fast reconstruction methods by using the FFT algorithm. The channel coding matrices are based on the finite Galois field algebra, which restricts the number of rows in such matrices to some subsets of the integer numbers. We also introduce means to combine these matrices to obtain sampling matrices with arbitrary number of rows.

Non-linear sampling methods are discussed in this thesis for the first time. When the sparsity domain is unknown at the time of sampling, no linear sampling method can guarantee perfect recovery; however, we show that non-linear methods can be used to recover $\backslash$-sparse signals. Furthermore, if the sparsity domain is known, non-linear methods can reduce both the number of required samples and the reconstruction complexity. The drawback of these methods is their sensitivity to additive noise.

Sparsity and compressibility are fundamental concepts in the field of compressed sensing. Although it is straightforward to define these concepts for finite dimensional vectors, the generalization to the infinite dimension and continuous domain is completely different. On the other hand, in order to be able to apply compressed sensing results to the real world problems, we need to consider continuous signals. Here we show that sparsity and compressibility concepts can be generalized to infinite deterministic and random sequences. Although the generalization from discrete to continuous signals is the main goal in many research works, the well-known generalization deals with substituting the vectors with matrices. For the latter case, instead of the zero/non-zero status of the elements, sparsity is usually defined through the rank of the matrix. In the last part of this thesis, we show how low-rank matrices can be retrieved from their point-wise distorted versions.

KEYWORDS

1. Compressed Sensing.
2. Sparsity.
3. Linear Projection.
4. Nonlinear Sampling.
5. i.i.d. Sequence.
6. Compressibility.
7. Low-rank Matrix.

SHARIF UNIVERSITY OF TECHNOLOGY
ELECTRICAL ENGINEERING DEPARTMENT

Ph.D. THESIS

Title:

Deterministic Compressed Sensing

by:

Arash Amini

Supervisor:
Prof. Farokh Marvasti

February 2011