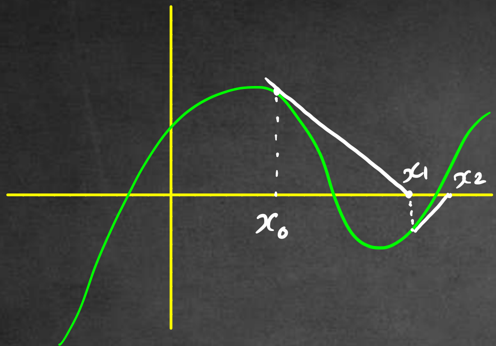
کنترل بهینه

جلسه چهارم ۹۹/۷/۷

روش‌های عددی حل مسائل بهینه‌سازی .



روش نیوتن

$$\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

$$\Phi(x) = 0$$

دنباله بازگشتی

$$y = \Phi(x_k) + D\Phi(x_k)(x - x_k) = 0$$

$$x_{k+1} = x_k - \left(D\Phi(x_k)\right)^{-1} \Phi(x_k)$$

اگر $\Phi(\bar{x})=0$ و $D\Phi(\bar{x})$ وارون‌پذیر باشد. به شرط آنکه x_0 به اندازه کافی نزدیک $\bar{x}$ انتخاب شود

آنگاه دنباله $\{x_k\}$ که به $\bar{x}$ همگرا است. به علاوه نرخ همگرایی حداقل درجه دو است .

تعریف- دنباله x_k به $\bar{x}$ همگرا با نرخ $p>0$ است، هرگاه p بزرگترین عددی باشد که

$$\lim_{k\to\infty} \frac{\|x^{k+1}-\bar{x}\|}{\|x^k-\bar{x}\|^p} = \beta < \infty$$

در روش نیوتن داریم:

$$x_{k+1}-\bar{x} = (x_k-\bar{x}) - D\phi(x_k)^{-1}\,\phi(x_k)$$

$$D\phi(x_k)\,(x_{k+1}-\bar{x}) = D\phi(x_k)\,(x_k-\bar{x}) - \phi(x_k)$$

$$0=\phi(\bar{x}) = \phi(x_k) + D\phi(x_k)(\bar{x}-x_k) + O(\|\bar{x}-x_k\|^2)$$

$$\Rightarrow \quad D\phi(x_k)(x_{k+1}-\bar{x}) = O(\|\bar{x}-x_k\|^2)$$

$$\|x_{k+1}-\bar{x}\| \leq C\,\|\bar{x}-x_k\|^2$$

نقاط ضعف : ۱ـ انتخاب نقطه شروع خیلی اثرگذار است.

۲ـ غیرتکینی بودن مشتق در فرایند تکرار مهم است.

۳ـ محاسبه وارون $D\Phi$ هزینه بر است. جدای از آن محاسبه مشتق Φ به خصوص زمانی که Φ پیچیده باشد سخت است.

مسأله بهینه سازی غیرمقید.

$$\min_{x \in \mathbb{R}^n} f(x)$$

به دنبال ریشه‌های $\nabla f(\bar{x}) = 0$ می‌گردیم. در عمل به یک روش نیوتن قرار می‌دهیم $\Phi = \nabla f$

$$\Phi = \nabla f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$D\Phi = H \rightsquigarrow f \text{ ماتریس هسیان}$$

$$d^k = x^{k+1} - x^k = -\left(H(x^k)\right)^{-1} \nabla f(x^k)$$

برای پیدا کردن نقطه مینیمم باید دقت کنیم که به نقاط بحرانی زینی یا ماکسیمم تابع f همگرا نشویم.

برای همین مناسب است که جهت d^k همیشه یک جهت کاهشی باشد. برای این منظور باید

$$\nabla f(x^k) \cdot d^k < 0$$

یا به طور معادل

$$\nabla f(x^k)^T \; H(x^k) \; \nabla f(x^k) > 0$$

که اگر $H(x^k)$ معین مثبت باشد، این شرط برقرار است.

استراتژی عمومی سازی : به دنبال برطرف کردن مشکل نقطه شروع الگوریتم هستیم.

برای این مطلب کافی است چک کنیم که در هر گام $f(x^{k+1}) < f(x^k)$

ایده جستجوی خطی : $d^k = -H(x^k)^{-1} \nabla f(x^k)$

$$x^{k+1} = x^k + \alpha d^k \qquad \alpha \geq 0$$

در این روش مقدار α را به گونه ای انتخاب می کنیم که

$$f(x^k + \alpha d^k) < f(x^k)$$

در عمل مثل این است که مسئله بهینه سازی در $\mathbb{R}^n$ را به یک مسئله بهینه سازی یک بعدی با قید $\alpha \geq 0$ تبدیل کرده ایم.

$$\min_{0 \leq \alpha} f(x^k + \alpha d^k)$$

قاعده Armijo : دو پارامتر $0<k_1<1$ و $1<k_2$ انتخاب کنید

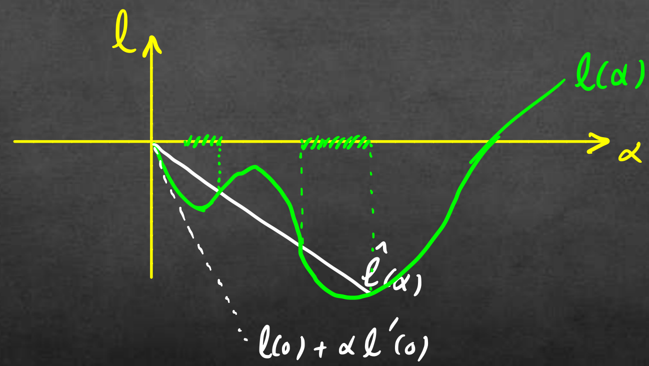
$$\ell(\alpha) := f(x^k + \alpha d^k)$$

$$\hat{\ell}(\alpha) := \ell(0) + k_1 \alpha \ell'(0)$$

$\bar{\alpha}$ انتخاب کنید به گونه‌ای که

$\ell(\bar{\alpha}) \le \hat{\ell}(\bar{\alpha})$ ← اجازه نمی‌دهد که طول گام خیلی بزرگ باشد

$\ell(k_2\bar{\alpha}) \ge \hat{\ell}(k_2\bar{\alpha})$ ← اجازه نمی‌دهد طول گام خیلی کوچک باشد



روش تصحیح بازگشتی : برای محاسبه ساده‌تر ماتریس هسیان به کار می‌رود.

$$d^k = -H(x^k)^{-1} \nabla f(x^k)$$

$$\nabla f(x^{k+1}) - \nabla f(x^k) \approx H(x^k) . (x^{k+1} - x^k)$$

$$[\gamma^0 | \dots | \gamma^{n-1}] \approx H(x^n) [\delta^0 | \dots | \delta^{n-1}]$$

$$\gamma^k = \nabla f(x^{k+1}) - \nabla f(x^k) \quad , \quad \delta^k = x^{k+1} - x^k$$

$$H(x^n)^{-1} \approx [\delta^0 | \dots | \delta^{n-1}] [\gamma^0 | \dots | \gamma^{n-1}]^{-1}$$

در این تخمین فقط از مشتقات مرتبه اول تابع استفاده شده است. اگر f یک تابع درجه ۲ باشد، آنگاه تقریب بالا دقیق است.

روش Broyden :

$$B^k = H(x^k)^{-1}$$

$$B^{k+1} = B^k + \frac{\delta^k {\delta^k}^T}{\gamma^k \cdot \delta^k} - \frac{B^k \gamma^k {\gamma^k}^T B^k}{{\gamma^k}^T B^k \gamma^k}$$

$$\delta^k = x^{k+1} - x^k \,, \quad \gamma^k = \nabla f(x^{k+1}) - \nabla f(x^k)$$

$$x^{k+1} = x^k + \alpha^k d^k$$

$$d^k = -B^k \nabla f(x^k)$$

نقطه شروع x^0 دلخواه و B^0 یک ماتریس معین مثبت دلخواه

تمرین -

$$f(x,y) = (1-x)^2 + 100(y-x^2)^2$$

$$\min_{\mathbb{R}^2} f \longrightarrow x = y = 1$$

به کمک دستور $fminuc$ در Matlab این مسئله را حل کنید.

$$fminuc(\,function\,,\, X_0\,,\, \underbrace{options}_{\downarrow\ optimset}\,)$$

نقطه شروع را $X_0 = (5,5)$ بگیرید و گام‌های رسیدن به جواب را رسم کنید.

بهینه سازی مقید

قید تساوی :

$$\min \ f(x)$$
$$\text{Subject to} \quad h(x)=0$$

$$h:\mathbb{R}^n \longrightarrow \mathbb{R}^m, \quad f:\mathbb{R}^n \longrightarrow \mathbb{R}$$

لاگرانژین $\mathcal{L}(x,\lambda) = f(x) + \lambda \cdot h(x)$

و مسأله زیر را حل میکنیم.

$$\begin{cases} \nabla_x \mathcal{L}(\bar{x},\bar{\lambda}) = 0 \\ \nabla_\lambda \mathcal{L}(\bar{x},\bar{\lambda}) = 0 \end{cases}$$

با روشهای شبه نیوتنی که گفته شد با قراردادن $\mathcal{L}$ به جای f مسأله را حل خواهیم کرد.

$$H = \begin{bmatrix} \nabla_{xx}\mathcal{L} & \nabla_{x\lambda}\mathcal{L} \\ \nabla_{\lambda x}\mathcal{L} & \nabla_{\lambda\lambda}\mathcal{L} \end{bmatrix} = \begin{bmatrix} \nabla_{xx}\mathcal{L} & Dh^T \\ Dh & 0 \end{bmatrix}$$

روش برنامه‌ریزی مربعی متوالی Successive quadratic programming (SQP)

در روش نیوتن

$$\begin{pmatrix} x^{k+1}-x^k \\ \lambda^{k+1}-\lambda^k \end{pmatrix} = -H \cdot \nabla L$$

$$\begin{pmatrix} d^k \\ \lambda^{k+1}-\lambda^k \end{pmatrix} = -\begin{pmatrix} D_{xx}L(x^k,\lambda^k) & Dh(x^k)^T \\ Dh(x^k) & 0 \end{pmatrix}\begin{pmatrix} \nabla f(x^k) + Dh(x^k)\lambda_k^T \\ h(x^k) \end{pmatrix}$$

بردار d^k که از تساوی بالا بدست می‌آید، جواب مسئله بهینه‌سازی درجه ۲ زیر است.

تقریب درجه ۲ تابع f $\rightarrow$ $$\min_{d} \; f(x^k) + \nabla f(x^k)\cdot d + \frac{1}{2} d^T \nabla_{xx} L(x^k,\lambda^k)\, d$$

تقریب خطی قید $h=0$ $\rightarrow$ Subject to $$h(x^k) + Dh(x^k)\, d = 0$$

بعد از حل مسئله بالا و پیدا کردن d^k و نقطه $x^{k+1}=x^k+d^k$، λ^{k+1} را از حل مسئله زیر پیدا می‌کنیم. $\nabla f(x^{k+1}) + Dh(x^{k+1})\,\lambda^{k+1}=0$

قیدیاسازی:

$$\min f(x)$$
$$\text{Subject to } g(x) \leq 0$$
$$h(x) = 0$$

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$h : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$
$$g : \mathbb{R}^n \longrightarrow \mathbb{R}^k$$

$$\mathcal{L}(x,\lambda,\nu) = f(x) + \nu \cdot g(x) + \lambda \cdot h(x) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^k_+ \longrightarrow \mathbb{R}$$

به صورت بازگشتی با داشتن (x^k, λ^k, ν^k) که $\nu^k \geq 0$ ، مسئله درجه ۲ زیر را حل می‌کنیم:

$$\min_{d} \; f(x^k) + \nabla f(x^k) \cdot d + \frac{1}{2} d^T D_{xx}\mathcal{L}(x^k, \lambda^k, \nu^k)\, d$$

$$\text{s.t.} \quad h(x^k) + Dh(x^k)\, d = 0$$

$$g(x^k) + Dg(x^k)\, d = 0$$

با پیدا کردن d^k و نقطه $x^{k+1} = x^k + d^k$ ، مقادیر λ^{k+1} و ν^{k+1} از حل دستگاه‌های زیر به دست می‌آیند:

$$\nabla f(x^{k+1}) + Dg(x^{k+1})^T \nu^{k+1} + Dh(x^{k+1})^T \lambda^{k+1} = 0$$

$$\underbrace{g(x^{k+1})}_{12} \cdot \nu^{k+1} = 0$$

$$[g(x^k) + Dg(x^k)\, d^k]^T \nu^{k+1} = 0$$