کنترل بهینه

جلسه ششم ۹۹/۷/۱۹

برنامه‌ریزی پویا :

$$x_{k+1} = f_k(x_k, u_k, w_k)$$

$$u_k \in U_k$$

$$J^*(x_0) = \min_{\substack{0 \le k \le n-1 \\ u_k \in U_k}} E\left\{ g_N(x_N) + \sum_{k=0}^{n-1} g_k(x_k, u_k, w_k) \right\}$$

قضیه : برای هر نقطه شروع x_0 ، تابع هزینه بهینه $J^*(x_0)$ برابر $J_0(x_0)$ است که J_0 به صورت بازگشتی از رابطه زیر محاسبه می‌شود :

$$J_N^*(x_N) = g_N(x_N)$$

$$J_k^*(x_k) = \min_{u_k \in U_k} E\left\{ g_k(x_k, u_k, w_k) + J_{k+1}^*\left(f_k(x_k, u_k, w_k)\right) \right\}$$

که امید ریاضی نسبت به توزیع احتمالی w_k محاسبه می‌شود. اگر $u_k = \mu_k^*(x_k)$ می‌نیمم سمت راست را بدهد $\pi^* = \{\mu_0^*, \dots, \mu_{N-1}^*\}$ سیاست بهینه است!

مثال - (مسابقه شطرنج) استراتژی بازی بزدلانه : با احتمال p_d مساوی می‌کند و با احتمال $1-p_d$ می‌بازد.

استراتژی بازی جسورانه : با احتمال p_w می‌برد و با احتمال $1-p_w$ می‌بازد.

$p_d > p_w$. در هر ست دو بازی انجام می‌گیرد و جمع امتیازات این دو بازی ملاک بردن آن ست است.

اگر بعد از دو بازی امتیازات شما بیشتر باشد، آن ست را برده‌اید و $+1$ امتیاز می‌گیرید.

اگر ست را ببازید، امتیازی نمی‌گیرید و رقیب $+1$ امتیاز می‌گیرد. اگر نتیجه دو بازی مساوی باشد، امتیاز شما p_w است.

سؤال: با چه استراتژی‌هایی بازی کنیم که بعد از N ست برنده باشیم؟

$x_k =$ اختلاف امتیاز شما با رقیب

$$J_N^*(x_N) = \begin{cases} 1 & x_N > 0 \\ P_\omega & x_N = 0 \\ 0 & x_N < 0 \end{cases}$$

$$J_K^*(x_k) = \max\Big[P_d\, J_{k+1}^*(x_k) + (1-P_d)\, J_{k+1}^*(x_k-1)\,,$$

$$P_\omega\, J_{k+1}^*(x_k+1) + (1-P_\omega)\, J_{k+1}^*(x_k-1) \Big]$$

$$x_{N-1} > 1 : \; J_{N-1}^*(x_{N-1}) = \max\big[P_d + (1-P_d)\,,\, P_\omega + (1-P_\omega) \big] = 1$$

استراتژی بازی مهم نیست.

$$x_{N-1} = 1 : \; J_{N-1}^*(x_{N-1}) = \max\big[P_d + P_\omega(1-P_d)\,,\; P_\omega + P_\omega(1-P_\omega) \big] = P_d + P_\omega(1-P_d)$$

استراتژی بهینه بزدلانه است.

$$J_{N-1}^*(0) = \max\big[P_d P_\omega\,,\, P_\omega \big] = P_\omega \longrightarrow$$ استراتژی جسورانه

$$J_{N-1}(-1) = p_w^2 \quad \rightarrow \quad \text{استراتژی بهینه جبرانی}$$

$$J_{N-1}(x_{N-1}) = 0 \qquad \text{if} \quad x_{N-1} < -1$$

در اکثر مسائل دینامیک مدل برحسب تابع $x_{k+1}=f_k(x_k,u_k,w_k)$ صریح داده شده است، مثل همین مثال بالا.

صورت‌بندی کلی مسئله بدین صورت است: اگر در حالت x_k باشید، مجاز هستید کنترل‌های $u_k \in U_k(x_k)$ را انتخاب کنید. برحسب انتخاب $u_k=u$ احتمال رفتن به حالت x_{k+1} با یک قانون احتمال داده شده است.

فرض کنید $x_k \in \{1,\dots,m\}$

$$P_{ij}(u) = P\left[x_{k+1}=j \mid x_k=i \,,\; u_k=u\right]$$

برای این مسائل می‌توانیم دینامیک f_k را به صورت زیر تعریف کنیم:

$$x_{k+1} = f_k(x_k,u_k,w_k) = w_k$$

$$P[\omega_k = j \mid x_k = i,\ u_k = u] = P_{ij}(u)$$

$$J_k^*(i) = \min_{u \in U(i)} E\left\{ g_k(i,u) + J_{k+1}^*(\omega_k) \right\}$$

$$= \min_{u \in U(i)} \left[g_k(i,u) + \sum_j P_{ij}(u)\, J_{k+1}^*(j) \right]$$

مثال ـ مسأله تعمیر ماشین (مثال ۷ از جلسه اول)

$$J_N^*(i) = 0 \qquad i = 1, \dots, n$$

$$J_k^*(i) = \min\left\{ R + g(1) + J_{k+1}^*(1),\ g(i) + \sum_{j \geq i} P_{ij}\, J_{k+1}(j) \right\}$$

$$J^*_{N-1}(i) = \min\{R + g(1),\ g(i)\}$$

$$g(1) \le g(2) \le \dots \le g(n)$$

کنترل بهینه این است که اگر حالت سیستم در وضعیتی است که $R+g(1) < g(i)$ ماشین را تعمیر کنند.

$$J^*_{N-2}(i) = \min\Big\{\underbrace{R + g(1) + J^*_{N-1}(1)}_{R+2g(1)},\ g(i) + \sum_{j \ge i} P_{ij} J_{N-1}(j)\Big\}$$

کنترل بهینه این است که اگر حالت سیستم از یه حدی وخیم‌تر بود، ماشین را باید تعمیر کرد.

مثال (مثال ۵ از جلسه اول)

$$J_N^*(i) = R(i) \qquad 1 \le i \le n$$

$$J_k^*(i) = \min \left\{ r(i) + C_f + \sum_j P_{ij}(u_f) J_{k+1}^*(j) \,,\; r(i) + C_s + \sum_j P_{ij}(u_s) J_{k+1}^*(j) \right\}$$

نکته - اگر دینامیک مسأله تأخیر زمانی (time lags) داشته باشد:

$$x_{k+1} = f_k(x_k, x_{k-1}, u_k, u_{k-1}, \omega_k)$$

$$x_1 = f_0(x_0, u_0, \omega_0)$$

متغیر جدید $y_k = x_{k-1}$ و $s_k = u_{k-1}$ را تعریف کنید:

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \\ s_{k+1} \end{pmatrix} = \begin{pmatrix} f_k(x_k, y_k, u_k, s_k, \omega_k) \\ x_k \\ u_k \end{pmatrix} = \tilde{f}_k(x_k, y_k, s_k, u_k, \omega_k)$$

که یک سیستم دینامیکی بدون تأخیر برای متغیر $\tilde{x}_k = (x_k, y_k, s_k)$ خواهد بود.

نکته- در مسأله‌های قبل متغیر تصادفی ω_k بدون حافظه است (یعنی ω_i و ω_j مستقل هستند)

اگر به عنوان مثال $\omega_{k+1} = \lambda \omega_k + \xi_k$ که ξ_k بدون حافظه است.

متغیر جدید $y_k = \omega_{k-1}$ تعریف کنید.

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} f_k(x_k, u_k, \lambda y_k + \xi_k) \\ \lambda y_k + \xi_k \end{pmatrix} = \tilde{f}_k(x_k, y_k, u_k, \xi_k)$$

معادله Hamilton-Jacobi-Bellman . (HJB)

$$(1)\begin{cases}\dot{x} = f(x,u,t)\\ x(t_0)=x_0\end{cases}$$

$$\min \leftarrow J = h(x(t_f),t_f) + \int_{t_0}^{t_f} g(x(t),u(t),t)\,dt$$

x_0 داده شده است و $u(t)\in U_{ad}$. نقطه پایانی t_f آزاد است یا از قید $m(x(t_f),t_f)=0$ محاسبه می‌شود.

در الگوریتم برنامه‌ریزی پویا، اگر در زمان t در نقطه y باشیم

$$J^*(y,t) = \min_u \left\{ \int_t^{t+\Delta t} g(x(s),u(s),s)\,ds + J^*(x(t+\Delta t),t+\Delta t) \right\}$$

$x(s)$ جواب (1) با شرط اولیه $x(t)=y$. subject to

$$J^*(x(t+\Delta t),\ t+\Delta t) = J^*(x(t),t) + \frac{\partial J^*}{\partial t}(x(t),t)\cdot\Delta t$$

$$+\ \nabla_x J^*(x(t),t)\cdot\big(x(t+\Delta t)-x(t)\big) + o(\Delta t)$$

با جایگذاری در رابطه قبل و با توجه به اینکه $x(t)=y$ است، خواهیم داشت:

$$\cancel{J^*(y,t)} = \min\left\{\int_t^{t+\Delta t} g(x(s),u(s),s)\,ds + \cancel{J^*(y,t)} + \frac{\partial J^*}{\partial t}(y,t)\,\Delta t\right.$$

$$\left. +\ \nabla_x J^*(y,t)\cdot\big(x(t+\Delta t)-x(t)\big) + o(\Delta t)\right\}$$

با تقسیم بر Δt و $\Delta t \to 0$ نتیجه می‌شود:

$$0 = \min_u\left\{ g(x(t),u(t),t) + \frac{\partial J^*}{\partial t}(y,t) + \nabla_x J^*(y,t)\cdot\dot{x}(t)\right\}$$

$$-\frac{\partial J^*}{\partial t}(y,t) = \min_u\left\{ g(y,u,t) + \nabla_x J^*(y,t)\cdot f(y,u,t)\right\}$$

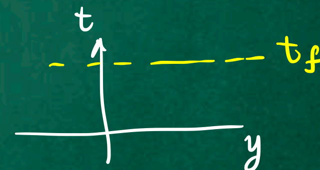
$$\mathcal{H}(y,u,\underbrace{\nabla_x J^*}_{\text{مجهول}},t) = g(y,u,t) + \nabla_x J^*(y,t)\cdot f(y,u,t) \leftarrow \text{Hamiltonian}$$

معادله (HJB):

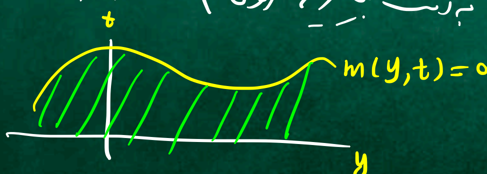
$$-\frac{\partial J^*}{\partial t} = \min_{u\in \mathcal{U}_{ad}} \mathcal{H}(y,u,\nabla_x J^*,t)$$

شرط مرزی این معادله با توجه به هزینه h روی زمان نهایی t_f به دست می‌آید. اگر t_f یک مقدار مشخص باشد:

$$J^*(y,t_f) = h(y,t_f)$$



اگر t_f آزاد بوده و از قید $m(x(t_f),t_f)=0$ به دست بیاید، روی منحنی $m(y,t)=0$ مقدار J^* برابر h خواهد بود.



حل -

$$\dot{x} = Ax + u$$

$$J = \frac{1}{2}|x(T)|^2 + \frac{1}{2}\int_0^T |u(t)|^2\,dt$$

$$\mathcal{H} = \frac{1}{2}|u|^2 + \nabla_x J^* \cdot (Ax + u)$$

اگر روی کنترل u قیدی نداشته باشیم، $\min_u \mathcal{H}$ در نقطه بحرانی $\frac{\partial \mathcal{H}}{\partial u} = 0$ بدست می‌آید.

$$\frac{\partial \mathcal{H}}{\partial u} = u + \nabla_x J^* \rightsquigarrow u = -\nabla_x J^*$$

مینیمم در نقطه $u = -\nabla_x J^*$ بدست می‌آید. $\left(\frac{\partial^2 \mathcal{H}}{\partial u^2} > 0\right)$

$$(\mathrm{HJB}) \begin{cases} -\frac{\partial J^*}{\partial t}(x,t) = -\frac{1}{2}|\nabla_x J^*(x,t)|^2 + \nabla_x J^*(x,t) \cdot Ax \\ J^*(x,T) = \frac{1}{2}|x|^2 \end{cases}$$

یک حدس معقول برای جواب (HJB) تابع $J(x,t)=\frac{1}{2}x^T B(t)\, x$ است که $B(t)$ از جایگذاری

در معادله پیدا می‌شود.

$$\begin{cases} -\frac{1}{2}x^T \dot{B}(t)\, x = -\frac{1}{2}\left| B(t)\, x \right|^2 + B(t)\, x \cdot A x \\ B(T) = I \end{cases}$$

$$\dot{B} = B^T B + 2A^T B$$