Aula 2

26 Fev 2019

Resumo da aula passada

- Conceitos preliminares: Horizontes de predição e controle, horizonte retrocedente.
- Elementos básicos de uma formulação de controle preditivo: Função de custo, modelo de predição (dinâmica da planta e perturbações), restrições.
- Exemplo: Dynamic Matrix Control (DMC)

Recapitulando: Notação adotada

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}(k+1|k) \\ \hat{y}(k+2|k) \\ \vdots \\ \hat{y}(k+N|k) \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} y_{ref} \\ y_{ref} \\ \vdots \\ y_{ref} \end{bmatrix} \triangleq [y_{ref}]_N$$

$$\Delta\hat{\mathbf{u}} = \begin{bmatrix} \Delta\hat{u}(k|k) \\ \Delta\hat{u}(k+1|k) \\ \vdots \\ \Delta\hat{u}(k+M-1|k) \end{bmatrix}$$

Recapitulando: Função de custo adotada no DMC

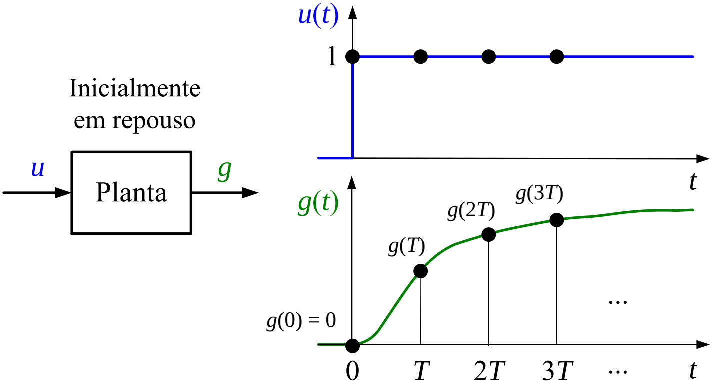
$$J(\hat{\mathbf{y}}, \Delta\hat{\mathbf{u}}) = (\hat{\mathbf{y}} - \mathbf{r})^T(\hat{\mathbf{y}} - \mathbf{r}) + \rho\Delta\hat{\mathbf{u}}^T\Delta\hat{\mathbf{u}}$$

- $\hat{\mathbf{y}}$ e $\Delta\hat{\mathbf{u}}$ não são independentes.
- A relação entre $\hat{\mathbf{y}}$ e $\Delta\hat{\mathbf{u}}$ é dada pelo modelo de predição adotado.
- Como se verá agora, assumindo um modelo de predição linear e invariante no tempo, a relação entre $\hat{\mathbf{y}}$ e $\Delta\hat{\mathbf{u}}$ toma a forma de uma **equação de predição** linear, que pode ser obtida com base na resposta a degrau da planta.

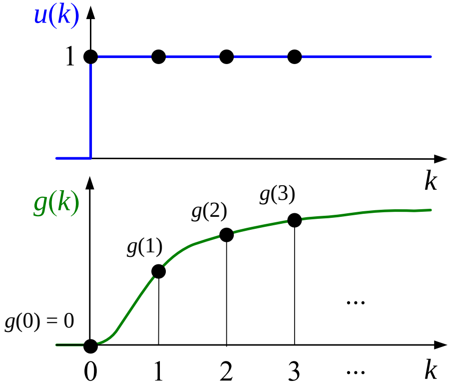
Resposta a degrau

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}, \text{ Sistema inicialmente em repouso}$$

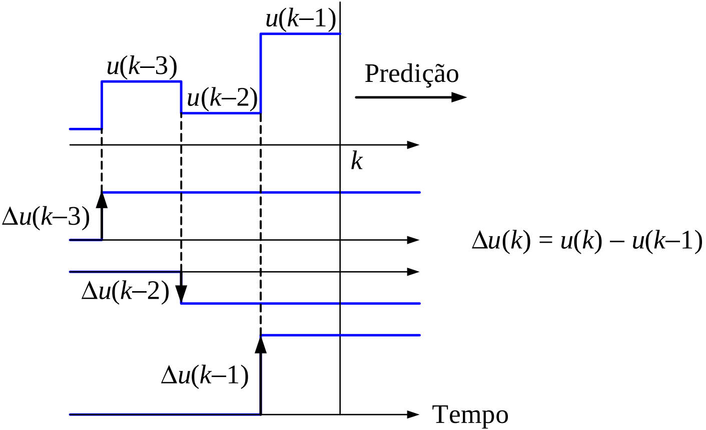
$$\Longrightarrow y(t) = g(t) \equiv \text{Resposta a Degrau.}$$



Seguindo a notação inicialmente convencionada, os valores de u e g no k-ésimo instante de amostragem serão indicados por $u(k)$ e $g(k)$, respectivamente.



Modelo baseado na resposta a degrau

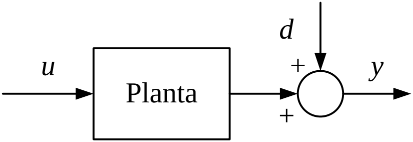


$$y(k) = g(1)\Delta u(k-1) + g(2)\Delta u(k-2) + \cdots = \sum_{n=1}^{\infty} g(n)\Delta u(k-n)$$

(Somatória de Convolução)

Estimação de perturbações

A formulação DMC supõe a existência de uma perturbação constante d na saída da planta:



$$y(k) = \sum_{n=1}^{\infty} g(n)\Delta u(k-n) + d(k)$$

$$y(k) = \sum_{n=1}^{\infty} g(n)\Delta u(k-n) + d(k)$$

Supondo que $y(k)$ seja medido por um sensor, pode-se obter uma estimativa para a perturbação $d(k)$ como sendo

$$d(k|k) = y(k) - \sum_{n=1}^{\infty} g(n)\Delta u(k-n)$$

Não está sendo usado chapéu em $d(k|k)$ por não se tratar de uma variável de otimização (de acordo com a notação adotada neste curso).

Predição um passo à frente

$$
\begin{aligned}
y(k) &= \sum_{n=1}^{\infty} g(n)\Delta u(k-n) + d(k) \\
y(k+1) &= \sum_{n=1}^{\infty} g(n)\Delta u(k+1-n) + d(k+1) = \\
&= g(1)\Delta u(k) + \sum_{n=2}^{\infty} g(n)\Delta u(k+1-n) + d(k+1)
\end{aligned}
$$

Em termos de valores preditos:

$$
\hat{y}(k+1|k) = g(1)\Delta\hat{u}(k|k) + \sum_{n=2}^{\infty} g(n)\Delta u(k+1-n) + d(k+1|k)
$$

Sob a hipótese de perturbação constante:

$$d(k+1|k) = d(k|k) = y(k) - \sum_{n=1}^{\infty} g(n)\Delta u(k-n)$$

Logo:

$$\begin{aligned}
\hat{y}(k+1|k) &= g(1)\Delta\hat{u}(k|k) + \sum_{n=2}^{\infty} g(n)\Delta u(k+1-n) + d(k+1|k) \\
&= g(1)\Delta\hat{u}(k|k) + \sum_{n=2}^{\infty} g(n)\Delta u(k+1-n) \\
&\quad + y(k) - \sum_{n=1}^{\infty} g(n)\Delta u(k-n)
\end{aligned}$$

$$
\begin{aligned}
\hat{y}(k+1|k) &= g(1)\Delta\hat{u}(k|k) + \overbrace{\sum_{n=2}^{\infty} g(n)\Delta u(k+1-n)}^{\sum_{n=1}^{\infty} g(n+1)\Delta u(k-n)} \\
&\quad + y(k) - \sum_{n=1}^{\infty} g(n)\Delta u(k-n)
\end{aligned}
$$

$$
\hat{y}(k+1|k) = g(1)\Delta\hat{u}(k|k) + y(k) + \sum_{n=1}^{\infty}[g(n+1) - g(n)]\Delta u(k-n)
$$

$$\hat{y}(k+1|k) = g(1)\Delta\hat{u}(k|k) + \underbrace{y(k) + \sum_{n=1}^{\infty}[g(n+1) - g(n)]\Delta u(k-n)}_{f(k+1|k)}$$

O termo

$$f(k+1|k) = y(k) + \sum_{n=1}^{\infty}[g(n+1) - g(n)]\Delta u(k-n)$$

é denominado "resposta livre" (*free response*) da planta.

$$\hat{y}(k+1|k) = g(1)\Delta\hat{u}(k|k) + f(k+1|k)$$

Cálculo da resposta livre

$$f(k+1|k) = y(k) + \sum_{n=1}^{\infty}[g(n+1) - g(n)]\Delta u(k-n)$$

Se a planta for estável, tem-se

$$g(n+1) - g(n) \overset{n\to\infty}{\longrightarrow} 0$$

Logo, a somatória na expressão de $f(k+1|k)$ pode ser truncada em um número N_s suficientemente grande de termos:

$$f(k+1|k) \simeq y(k) + \sum_{n=1}^{N_s-1}[g(n+1) - g(n)]\Delta u(k-n)$$

(assumindo $g(n) \simeq g(N_s), \forall n \geq N_s$).

Predição i passos à frente

$$
\begin{aligned}
y(k) &= \sum_{n=1}^{\infty} g(n)\Delta u(k-n) + d(k) \\
y(k+i) &= \sum_{n=1}^{\infty} g(n)\Delta u(k+i-n) + d(k+i) = \\
&= \sum_{n=1}^{i} g(n)\Delta u(k+i-n) + \sum_{n=i+1}^{\infty} g(n)\Delta u(k+i-n) + d(k+i)
\end{aligned}
$$

Em termos de valores preditos:

$$
\begin{aligned}
\hat{y}(k+i|k) &= \sum_{n=1}^{i} g(n)\Delta \hat{u}(k+i-n|k) + \sum_{n=i+1}^{\infty} g(n)\Delta u(k+i-n) \\
&\quad + d(k+i|k)
\end{aligned}
$$

Sob a hipótese de perturbação constante:

$$d(k+i|k) = d(k|k) = y(k) - \sum_{n=1}^{\infty} g(n)\Delta u(k-n)$$

Logo:

$$\begin{aligned}\hat{y}(k+i|k) \;&=\; \sum_{n=1}^{i} g(n)\Delta\hat{u}(k+i-n|k) \;+\; \sum_{n=i+1}^{\infty} g(n)\Delta u(k+i-n) \\ &+\; y(k) - \sum_{n=1}^{\infty} g(n)\Delta u(k-n)\end{aligned}$$

$$\hat{y}(k+i|k) = \sum_{n=1}^{i} g(n)\Delta\hat{u}(k+i-n|k) + \overbrace{\sum_{n=i+1}^{\infty} g(n)\Delta u(k+i-n)}^{\sum_{n=1}^{\infty} g(n+i)\Delta u(k-n)}$$
$$+ \; y(k) - \sum_{n=1}^{\infty} g(n)\Delta u(k-n)$$

$$\hat{y}(k+i|k) = \sum_{n=1}^{i} g(n)\Delta\hat{u}(k+i-n|k)$$
$$+ \; \underbrace{y(k) + \sum_{n=1}^{\infty}[g(n+i)-g(n)]\Delta u(k-n)}_{f(k+i|k)}$$

$$\hat{y}(k+i|k) = \sum_{n=1}^{i} g(n)\Delta\hat{u}(k+i-n|k) + f(k+i|k)$$

em que

$$f(k+i|k) \simeq y(k) + \sum_{n=1}^{N_s-1} [g(n+i) - g(n)]\Delta u(k-n)$$

sendo N_s suficientemente grande para que $g(N_s+i) \simeq g(N_s)$.

Equação de predição: Uso de notação matricial

$$\hat{y}(k+i|k) \;=\; \sum_{n=1}^{i} g(n)\Delta\hat{u}(k+i-n|k) + f(k+i|k)$$

$$\begin{aligned}
\hat{y}(k+1|k) &= g(1)\Delta\hat{u}(k|k) + f(k+1|k) \\
\hat{y}(k+2|k) &= g(2)\Delta\hat{u}(k|k) + g(1)\Delta\hat{u}(k+1|k) + f(k+2|k) \\
&\vdots \\
\hat{y}(k+N|k) &= g(N)\Delta\hat{u}(k|k) + g(N-1)\Delta\hat{u}(k+1|k) + \cdots \\
&\quad + g(1)\Delta\hat{u}(k+N-1|k) + f(k+N|k)
\end{aligned}$$

$$
\begin{bmatrix} \hat{y}(k+1|k) \\ \hat{y}(k+2|k) \\ \vdots \\ \hat{y}(k+N|k) \end{bmatrix} = \begin{bmatrix} g(1) & 0 & \cdots & 0 \\ g(2) & g(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g(N) & g(N-1) & \cdots & g(1) \end{bmatrix} \begin{bmatrix} \Delta\hat{u}(k|k) \\ \Delta\hat{u}(k+1|k) \\ \vdots \\ \Delta\hat{u}(k+N-1|k) \end{bmatrix}
$$

$$
+ \begin{bmatrix} f(k+1|k) \\ f(k+2|k) \\ \vdots \\ f(k+N|k) \end{bmatrix}
$$

$$\overbrace{\begin{bmatrix} \hat{y}(k+1|k) \\ \hat{y}(k+2|k) \\ \vdots \\ \hat{y}(k+N|k) \end{bmatrix}}^{\hat{\mathbf{y}}(N\times 1)} = \overbrace{\begin{bmatrix} g(1) & 0 & \cdots & 0 \\ g(2) & g(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g(N) & g(N-1) & \cdots & g(1) \end{bmatrix}}^{G(N\times N)} \overbrace{\begin{bmatrix} \Delta\hat{u}(k|k) \\ \Delta\hat{u}(k+1|k) \\ \vdots \\ \Delta\hat{u}(k+N-1|k) \end{bmatrix}}^{\Delta\hat{\mathbf{u}}(N\times 1)}$$

$$+ \underbrace{\begin{bmatrix} f(k+1|k) \\ f(k+2|k) \\ \vdots \\ f(k+N|k) \end{bmatrix}}_{\mathbf{f}(N\times 1)} \Longrightarrow \boxed{\hat{\mathbf{y}} = G\Delta\hat{\mathbf{u}} + \mathbf{f}}$$

G: “Matriz Dinâmica” (formato *Toeplitz*).

Equação de predição: Caso $M < N$

Se o horizonte de controle M for menor que o horizonte de predição N, deve-se impor $\Delta\hat{u}(k+i-1|k) = 0$, $M+1 \leq i \leq N$. Logo:

$$\hat{\mathbf{y}} = \left[\begin{array}{cccc|ccc} g(1) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ g(2) & g(1) & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g(M) & g(M-1) & \cdots & g(1) & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g(N-1) & g(N-2) & \cdots & g(N-M) & g(N-M-1) & \cdots & 0 \\ g(N) & g(N-1) & \cdots & g(N-M+1) & g(N-M) & \cdots & g(1) \end{array}\right]$$

$$\times$$

$$[\Delta\hat{u}(k|k)\ \Delta\hat{u}(k+1|k)\ \cdots\ \Delta\hat{u}(k+M-1|k)\ |\ 0 \cdots 0]^T + \mathbf{f}$$

$$\hat{\mathbf{y}} = \overbrace{\begin{bmatrix} g(1) & 0 & \cdots & 0 \\ g(2) & g(1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ g(M) & g(M-1) & \cdots & g(1) \\ \vdots & \vdots & \vdots & \vdots \\ g(N-1) & g(N-2) & \cdots & g(N-M) \\ g(N) & g(N-1) & \cdots & g(N-M+1) \end{bmatrix}}^{G(N\times M)}$$

$$\times \overbrace{[\Delta\hat{u}(k|k)\ \Delta\hat{u}(k+1|k)\ \cdots\ \Delta\hat{u}(k+M-1|k)]^T}^{\Delta\hat{\mathbf{u}}(M\times 1)} + \mathbf{f}$$

$$\Longrightarrow \boxed{\hat{\mathbf{y}} = G\Delta\hat{\mathbf{u}} + \mathbf{f}}$$

Solução do problema de otimização

Problema: Minimizar

$$J(\hat{\mathbf{y}}, \Delta\hat{\mathbf{u}}) = (\hat{\mathbf{y}} - \mathbf{r})^T(\hat{\mathbf{y}} - \mathbf{r}) + \rho\Delta\hat{\mathbf{u}}^T\Delta\hat{\mathbf{u}}$$

s.a.

$$\hat{\mathbf{y}} = G\Delta\hat{\mathbf{u}} + \mathbf{f}$$

com respeito a $\hat{\mathbf{y}} \in \mathbb{R}^N$, $\Delta\hat{\mathbf{u}} \in \mathbb{R}^M$.

$$J(\hat{\mathbf{y}}, \Delta\hat{\mathbf{u}}) = (\hat{\mathbf{y}} - \mathbf{r})^T(\hat{\mathbf{y}} - \mathbf{r}) + \rho\Delta\hat{\mathbf{u}}^T\Delta\hat{\mathbf{u}} \tag{1}$$

$$\hat{\mathbf{y}} = G\Delta\hat{\mathbf{u}} + \mathbf{f} \tag{2}$$

Substituindo a expressão de $\hat{\mathbf{y}}$ dada em (2) na função de custo (1), chega-se a

$$\begin{aligned} J &= (G\Delta\hat{\mathbf{u}} + \mathbf{f} - \mathbf{r})^T(G\Delta\hat{\mathbf{u}} + \mathbf{f} - \mathbf{r}) + \rho\Delta\hat{\mathbf{u}}^T\Delta\hat{\mathbf{u}} = \\ &= \Delta\hat{\mathbf{u}}^T G^T G\Delta\hat{\mathbf{u}} + \overbrace{(G\Delta\hat{\mathbf{u}})^T(\mathbf{f} - \mathbf{r}) + (\mathbf{f} - \mathbf{r})^T(G\Delta\hat{\mathbf{u}})}^{2(\mathbf{f}-\mathbf{r})^T(G\Delta\hat{\mathbf{u}})} + \\ &+ (\mathbf{f} - \mathbf{r})^T(\mathbf{f} - \mathbf{r}) + \rho\Delta\hat{\mathbf{u}}^T\Delta\hat{\mathbf{u}} \end{aligned}$$

$$
\begin{aligned}
J &= \Delta\hat{\mathbf{u}}^T G^T G \Delta\hat{\mathbf{u}} + \overbrace{(G\Delta\hat{\mathbf{u}})^T(\mathbf{f}-\mathbf{r}) + (\mathbf{f}-\mathbf{r})^T(G\Delta\hat{\mathbf{u}})}^{2(\mathbf{f}-\mathbf{r})^T(G\Delta\hat{\mathbf{u}})} + \\
&+ (\mathbf{f}-\mathbf{r})^T(\mathbf{f}-\mathbf{r}) + \rho\Delta\hat{\mathbf{u}}^T\Delta\hat{\mathbf{u}}
\end{aligned}
$$

$$
= \Delta\hat{\mathbf{u}}^T \underbrace{(G^T G + \rho I)}_{(1/2)\mathcal{H}} \Delta\hat{\mathbf{u}} + \underbrace{2(\mathbf{f}-\mathbf{r})^T G}_{c^T} \Delta\hat{\mathbf{u}} + \underbrace{(\mathbf{f}-\mathbf{r})^T(\mathbf{f}-\mathbf{r})}_{\mathrm{cte}}
$$

$$
J = \frac{1}{2}\Delta\hat{\mathbf{u}}^T \mathcal{H} \Delta\hat{\mathbf{u}} + c^T\Delta\hat{\mathbf{u}} + \mathrm{cte}
$$

sendo

$$
\mathcal{H} = 2(G^T G + \rho I),\ c = 2G^T(\mathbf{f}-\mathbf{r}),\ \mathrm{cte} = (\mathbf{f}-\mathbf{r})^T(\mathbf{f}-\mathbf{r})
$$

Problema reformulado

Minimizar

$$J(\Delta\hat{\mathbf{u}}) = \frac{1}{2}\Delta\hat{\mathbf{u}}^T \mathcal{H} \Delta\hat{\mathbf{u}} + c^T \Delta\hat{\mathbf{u}} + \mathrm{cte}$$

com respeito a $\Delta\hat{\mathbf{u}} \in \mathbb{R}^M$.

Problema de otimização sem restrições

Revisão de Otimização sem Restrições

Mini-Tutorial: Matrizes Positivo/Negativo-Definidas

Uma matriz $Q \in \mathbb{R}^{n \times n}$ simétrica é dita positivo-definida (PD) se e somente se (s.s.s)

$$x^T Q x > 0, \forall x \in \mathbb{R}^n, x \neq 0$$

- Notações comumente empregadas: $Q = Q^T > 0$, $Q = Q^T \succ 0$.
- Termo alternativo: Definida positiva

Relação com os autovalores de Q

Sejam $\lambda_1, \lambda_2, \ldots, \lambda_n$ os autovalores de uma matriz $Q = Q^T \in \mathbb{R}^{n \times n}$.

Tem-se que $Q > 0 \iff \lambda_i > 0,\ i = 1, 2, \ldots, n$.

$$\boxed{Q > 0 \Longleftrightarrow \lambda_i > 0,\ i = 1, 2, \ldots, n}$$

Demonstração: Sejam $v_1, v_2, \ldots, v_n$ os autovetores de Q (com norma unitária) associados aos autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$. Tendo em vista que $Qv_i = \lambda_i v_i,\ i = 1, 2, \ldots, n$, pode-se escrever

$$Q \left[v_1 \ v_2 \ \cdots \ v_n \right] = \underbrace{\left[v_1 \ v_2 \ \cdots \ v_n \right]}_{V} \underbrace{\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}}_{\Lambda}$$

ou seja $QV = V\Lambda$.

Como $Q = Q^T \in \mathbb{R}^{n \times n}$, sabe-se que os autovalores $\lambda_1, \lambda_2, \ldots, \lambda_n$ são reais e os autovetores $v_1, v_2, \ldots, v_n$ são mutuamente ortogonais.

Com efeito, se λ é um autovalor de Q associado a um autovetor v, tem-se

$$Qv = \lambda v$$

Multiplicando os dois lados dessa identidade por $v^\star$, em que $\star$ denota o complexo-conjugado transposto, obtém-se

$$v^\star Q v = v^\star \lambda v = \lambda v^\star v$$

Por outro lado, tem-se que $v^\star Q v \in \mathbb{R}$, pois $(v^\star Q v)^\star = v^\star Q^\star v = v^\star Q v$. Portanto, como $v^\star v \in \mathbb{R}$, deve-se ter $\lambda \in \mathbb{R}$.

Uma prova simples da ortogonalidade entre v_i e v_j, $i \neq j$, pode ser construída para o caso particular em que $\lambda_i \neq \lambda_j$.

Neste caso, pode-se escrever

$$Qv_i = \lambda_i v_i \quad (3)$$

$$Qv_j = \lambda_j v_j \quad (4)$$

Pré-multiplicando (3) e (4) por v_j^T e v_i^T, respectivamente, tem-se

$$v_j^T Q v_i = \lambda_i v_j^T v_i \quad (5)$$

$$v_i^T Q v_j = \lambda_j v_i^T v_j \quad (6)$$

Subtraindo (6) de (5) chega-se a $(\lambda_i - \lambda_j) v_i^T v_j = 0$. Como $\lambda_i \neq \lambda_j$, por hipótese, conclui-se que $v_i^T v_j = 0$.

Para o caso de autovalores com multiplicidade maior do que um, ver GANTMACHER, F. R. *The Theory of Matrices*, 2 ed. New York: Chelsea, 1990 (v. 1, p. 270-272).

Retornando à demonstração principal:

$$QV = V\Lambda \tag{7}$$

Como $v_1, v_2, \ldots, v_n$ possuem norma unitária e são mutuamente ortogonais, tem-se que

$$V^T V = VV^T = I$$

sendo I uma matriz identidade.

Logo, pós-multiplicando os dois lados de (7) por V^T, tem-se que

$$QVV^T = V\Lambda V^T \Rightarrow Q = V\Lambda V^T \tag{8}$$

sendo

$$V = \left[v_1 \; v_2 \; \cdots \; v_n \right], \; \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{9}$$

$$Q = V\Lambda V^T, \; V = [v_1 \; v_2 \; \cdots \; v_n], \; \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Portanto, dado $x \in \mathbb{R}^n$ tem-se que

$$x^T Q x = x^T V \Lambda V^T x = z^T \Lambda z$$

sendo $z = V^T x$. Por outro lado,

$$z^T \Lambda z = [z_1 \; z_2 \; \cdots \; z_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

$$= \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2$$

Demonstração da suficiência:

- Hipótese: $\lambda_i > 0, i = 1, 2, \ldots, n$.
- Tese: $Q > 0$ (isto é, $x \neq 0 \Rightarrow x^T Q x > 0$)

Seja $x \neq 0$. Tomando $z = V^T x$, tem-se que

$$||z||^2 = z^T z = x^T V V^T x = x^T x \neq 0$$

e portanto $z \neq 0$.

Sabe-se ainda que

$$x^T Q x = z^T \Lambda z = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2$$

Como $z \neq 0$, tem-se que $z_i \neq 0$ para algum i. Como $\lambda_i > 0$, por hipótese, segue que

$$x^T Q x = z^T \Lambda z \geq \lambda_i z_i^2 > 0$$

Portanto, $x^T Q x > 0$, cqd.

Demonstração da necessidade:

- Hipótese: $Q > 0$ (isto é, $x \neq 0 \Rightarrow x^T Q x > 0$)
- Tese: $\lambda_i > 0, i = 1, 2, \ldots, n$.

Por absurdo, suponha que $\lambda_i \leq 0$ para algum i. Seja ainda $z = [0 \ \cdots \ 0 \ z_i \ 0 \cdots 0]^T$ com $z_i \neq 0$ e tome-se $x = Vz$. Tem-se então que

$$||x||^2 = x^T x = z^T V^T V z = z^T z = z_i^2 > 0$$

e, portanto, $x \neq 0$. Por outro lado,

$$x^T Q x = z^T \Lambda z = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2 = \lambda_i z_i^2 \leq 0$$

dado que $\lambda_i \leq 0$, por hipótese. Verifica-se então que é possível ter $x^T Q x \leq 0$ com $x \neq 0$, o que contradiz a hipótese de que $Q > 0$. Portanto, assumir que $\lambda_i \leq 0$ para algum i conduz a uma contradição. Logo, todos os autovalores $\lambda_i, i = 1, 2, \ldots, n$, devem ser estritamente positivos, c.q.d.

Observações:

- Uma matriz PD pode ter elementos negativos.

  Por exemplo, a matriz

  $$Q = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$$

  tem autovalores $\lambda_1 = 1{,}4$ e $\lambda_2 = 3{,}6$.

- Uma matriz com todos os elementos positivos não é necessariamente PD.

  Por exemplo, a matriz

  $$Q = \begin{bmatrix} 2 & 4 \\ 4 & 3 \end{bmatrix}$$

  tem autovalores $\lambda_1 = -1{,}5$ e $\lambda_2 = 6{,}5$.

Definições adicionais

Seja $Q = Q^T \in \mathbb{R}^{n \times n}$. Diz-se que:

- $Q > 0$ (Positivo-Definida) se $x^T Qx > 0, \forall x \neq 0$.
- $Q \geq 0$ (Positivo-Semidefinida) se $x^T Qx \geq 0, \forall x$.
- $Q < 0$ (Negativo-Definida) se $x^T Qx < 0, \forall x \neq 0$.
- $Q \leq 0$ (Negativo-Semidefinida) se $x^T Qx \leq 0, \forall x$.
- Q é Indefinida nos demais casos.

Condições sobre os autovalores:

- $Q > 0 \Leftrightarrow \lambda_i > 0, i = 1, 2, \ldots, n$
- $Q \geq 0 \Leftrightarrow \lambda_i \geq 0, i = 1, 2, \ldots, n$
- $Q < 0 \Leftrightarrow \lambda_i < 0, i = 1, 2, \ldots, n$
- $Q \leq 0 \Leftrightarrow \lambda_i \leq 0, i = 1, 2, \ldots, n$
- Q indefinida $\Leftrightarrow \lambda_i > 0$ e $\lambda_j < 0$ para algum i e j.

Condições de otimalidade

Seja uma função $J : \mathbb{R}^n \to \mathbb{R}$. Um ponto $x^* \in \mathbb{R}^n$ é dito ser um mínimo local de J se $J(x^*) \leq J(x)$ para todo x em uma vizinhança de x^*.

Se $J(x^*) \leq J(x)$ para todo $x \in \mathbb{R}^n$, diz-se que x^* é um mínimo global de J.

Notações comumente empregadas:

$$x^* = \arg \min_{x \in \mathbb{R}^n} J(x)$$
$$J(x^*) = \min_{x \in \mathbb{R}^n} J(x)$$

O problema de otimização é expresso como

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

Teorema de Taylor - Caso univariado

Seja uma função $J : \mathbb{R} \to \mathbb{R}$. Se $J \in C^r$ (isto é, se J possui derivadas contínuas até ordem r), então dados $x \in \mathbb{R}$ e $h \in \mathbb{R}$, existe $\theta \in [0,1]$ tal que

$$J(x+h) = J(x) + hJ'(x) + \frac{1}{2}h^2 J''(x) + \cdots + \\ + \frac{1}{(r-1)!}h^{r-1}J^{(r-1)}(x) + h^r \frac{1}{r!}J^{(r)}(x+\theta h)$$

Em particular, se $J \in C^2$, então

$$J(x+h) = J(x) + hJ'(x) + \frac{1}{2}h^2 J''(x+\theta h)$$

Referência: GILL, P.E.; MURRAY, W.; WRIGHT, M.H. *Practical Optimization*, Academic Press, 1981.

Condições necessárias para otimalidade - Caso univariado

Seja uma função $J : \mathbb{R} \rightarrow \mathbb{R}$ pertencente à classe C^2. Se $x^* \in \mathbb{R}$ é um mínimo local de f, então as seguintes condições devem ser satisfeitas:

- $J'(x^*) = 0$
- $J''(x^*) \geq 0$

Condições suficientes para otimalidade - Caso univariado

Seja uma função $J : \mathbb{R} \rightarrow \mathbb{R}$ pertencente à classe C^2. Se as seguintes condições forem satisfeitas:

- $J'(x^*) = 0$
- $J''(x^*) > 0$

então $x^* \in \mathbb{R}$ é um mínimo local de J.

Teorema de Taylor - Caso multivariado com $r = 2$

Seja uma função $J : \mathbb{R}^n \to \mathbb{R}$ da classe C^2. Dados $x \in \mathbb{R}^n$ e $\Delta x \in \mathbb{R}^n$, existe $\theta \in [0, 1]$ tal que

$$J(x + \Delta x) = J(x) + \left[\frac{\partial J}{\partial x}(x)\right]^T \Delta x + \frac{1}{2}\Delta x^T \left[\frac{\partial^2 J}{\partial x^2}(x + \theta \Delta x)\right] \Delta x$$

em que $\frac{\partial J}{\partial x} \in \mathbb{R}^n$ (vetor gradiente) e $\frac{\partial^2 J}{\partial x^2} \in \mathbb{R}^{n \times n}$ (matriz Hessiana, ou "de curvatura") são dados por

$$\frac{\partial J}{\partial x} = \begin{bmatrix} \frac{\partial J}{\partial x_1} \\ \frac{\partial J}{\partial x_2} \\ \vdots \\ \frac{\partial J}{\partial x_n} \end{bmatrix}, \quad \frac{\partial^2 J}{\partial x^2} = \begin{bmatrix} \frac{\partial^2 J}{\partial x_1^2} & \frac{\partial^2 J}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 J}{\partial x_1 \partial x_n} \\ \frac{\partial^2 J}{\partial x_2 \partial x_1} & \frac{\partial^2 J}{\partial x_2^2} & \cdots & \frac{\partial^2 J}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 J}{\partial x_n \partial x_1} & \frac{\partial^2 J}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 J}{\partial x_n^2} \end{bmatrix}$$

Observação sobre a matriz Hessiana

$$\frac{\partial^2 J}{\partial x^2} = \begin{bmatrix} \frac{\partial^2 J}{\partial x_1^2} & \frac{\partial^2 J}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 J}{\partial x_1 \partial x_n} \\ \frac{\partial^2 J}{\partial x_2 \partial x_1} & \frac{\partial^2 J}{\partial x_2^2} & \cdots & \frac{\partial^2 J}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 J}{\partial x_n \partial x_1} & \frac{\partial^2 J}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 J}{\partial x_n^2} \end{bmatrix}$$

Como as derivadas de segunda ordem são contínuas (no caso $r = 2$), tem-se que

$$\frac{\partial^2 J}{\partial x_i \partial x_j} = \frac{\partial^2 J}{\partial x_j \partial x_i}$$

Portanto, a matriz Hessiana $\frac{\partial^2 J}{\partial x^2}$ é simétrica.

Condições necessárias para otimalidade - Caso multivariado

Seja uma função $J : \mathbb{R}^n \rightarrow \mathbb{R}$ pertencente à classe C^2. Se $x^* \in \mathbb{R}^n$ é um mínimo local de J, então as seguintes condições devem ser satisfeitas:

- $\dfrac{\partial J}{\partial x}(x^*) = 0$
- $\dfrac{\partial^2 J}{\partial x^2}(x^*) \geq 0$

Observação: Se $\dfrac{\partial J}{\partial x}(x^*) = 0$, diz-se que x^* é um "ponto estacionário" de J.

Condições suficientes para otimalidade - Caso multivariado

Seja uma função $J : \mathbb{R}^n \rightarrow \mathbb{R}$ pertencente à classe C^2. Se as seguintes condições forem satisfeitas:

- $\dfrac{\partial J}{\partial x}(x^*) = 0$
- $\dfrac{\partial^2 J}{\partial x^2}(x^*) > 0$

então $x^* \in \mathbb{R}^n$ é um mínimo local de J.

Observação: Suponha que $J \in C^2$ e $\dfrac{\partial J}{\partial x}(x^*) = 0$. Tem-se então:

- $\dfrac{\partial^2 J}{\partial x^2}(x^*) > 0 \Rightarrow x^*$ é mínimo local.
- $\dfrac{\partial^2 J}{\partial x^2}(x^*) < 0 \Rightarrow x^*$ é máximo local.
- $\dfrac{\partial^2 J}{\partial x^2}(x^*)$ indefinida $\Rightarrow x^*$ é ponto de sela.

Algumas expressões para cálculo de gradientes

Sejam $x \in \mathbb{R}^n, y \in \mathbb{R}^n, Q \in \mathbb{R}^{n \times n}$. Então:

$$\frac{\partial(y^T x)}{\partial x} = \frac{\partial(x^T y)}{\partial x} = y$$

$$\frac{\partial(y^T Q x)}{\partial x} = \frac{\partial(x^T Q^T y)}{\partial x} = Q^T y$$

$$\frac{\partial(x^T Q x)}{\partial x} = Qx + Q^T x$$

$$\frac{\partial\left[(x-y)^T Q (x-y)\right]}{\partial x} = (Q + Q^T)(x - y)$$

Se Q for simétrica, as expressões se simplificam, pois $Q + Q^T = 2Q$.

Algumas expressões para cálculo de Hessianas

Sejam $x \in \mathbb{R}^n, y \in \mathbb{R}^n, Q \in \mathbb{R}^{n \times n}$. Então:

$$\frac{\partial^2 (x^T Q x)}{\partial x^2} = Q + Q^T$$

$$\frac{\partial^2 \left[(x-y)^T Q (x-y) \right]}{\partial x^2} = Q + Q^T$$

Se Q for simétrica, as expressões se simplificam, pois $Q + Q^T = 2Q$.

Ex: Funções Quadráticas

$$J(x) = \frac{1}{2} x^T \mathcal{H} x + c^T x + \mathrm{cte}$$

Neste caso, tem-se

$$\frac{\partial J}{\partial x}(x) = \mathcal{H} x + c$$

Portanto, x^* deve satisfazer $\mathcal{H} x^* + c = 0$, isto é

$$x^* = -\mathcal{H}^{-1} c$$

desde que $\mathcal{H}$ seja não singular.

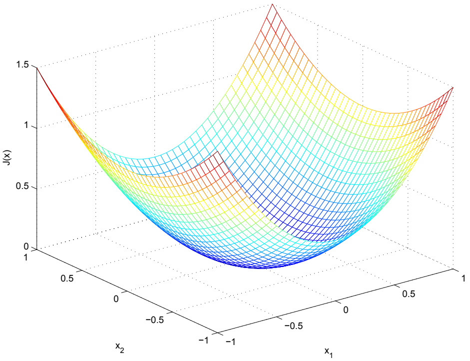
Adicionalmente, a matriz Hessiana em x^* é dada por

$$\frac{\partial^2 J}{\partial x^2}(x^*) = \mathcal{H}$$

Exemplo 1

$$x^* = -\mathcal{H}^{-1}c, \; \frac{\partial^2 J}{\partial x^2}(x^*) = \mathcal{H}$$

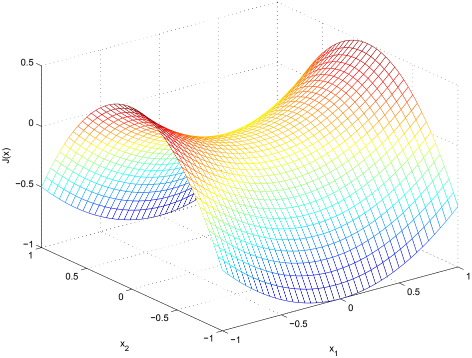
$$\mathcal{H} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \; c = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$



Exemplo 2

$$x^* = -\mathcal{H}^{-1}c, \; \frac{\partial^2 J}{\partial x^2}(x^*) = \mathcal{H}$$

$$\mathcal{H} = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}, \; c = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$



Retornando ao problema inicial (DMC)

Minimizar

$$J(\Delta\hat{\mathbf{u}}) = \frac{1}{2}\Delta\hat{\mathbf{u}}^T \mathcal{H} \Delta\hat{\mathbf{u}} + c^T \Delta\hat{\mathbf{u}} + \mathrm{cte}$$

com respeito a $\Delta\hat{\mathbf{u}} \in \mathbb{R}^M$, sendo

$$\mathcal{H} = 2(G^T G + \rho I),\ c = 2G^T(\mathbf{f} - \mathbf{r})$$

$$\rho > 0$$

Vale notar que a matrix Hessiana $\mathcal{H}$ é simétrica e positivo-definida. Com efeito, dado $x \neq 0$, tem-se que

$$x^T \mathcal{H} x = 2x^T(G^T G + \rho I)x = 2\underbrace{x^T(G^T G)x}_{\geq 0} + 2\rho\underbrace{x^T x}_{>0} > 0$$

$$J(\Delta\hat{\mathbf{u}}) = \frac{1}{2}\Delta\hat{\mathbf{u}}^T \mathcal{H} \Delta\hat{\mathbf{u}} + c^T \Delta\hat{\mathbf{u}} + \mathrm{cte}$$

$$\mathcal{H} > 0$$

Trata-se de uma função quadrática em $\Delta\hat{\mathbf{u}}$ com matriz Hessiana positivo-definida. Assim, como visto anteriormente,

$$\Delta\hat{\mathbf{u}}^* = -\mathcal{H}^{-1} c$$

com

$$\mathcal{H} = 2(G^T G + \rho I), \; c = 2G^T(\mathbf{f} - \mathbf{r})$$

Logo,

$$\Delta\hat{\mathbf{u}}^* = (G^T G + \rho I)^{-1} G^T (\mathbf{r} - \mathbf{f})$$

Observações

$$\Delta\hat{\mathbf{u}}^* = (G^T G + \rho I)^{-1} G^T (\mathbf{r} - \mathbf{f})$$

- A matriz $(G^T G + \rho I)$ tem dimensões $M \times M$
- O parâmetro $\rho > 0$ pode ser usado para ajustar o esforço de controle.
- No instante atual k, a variação de controle a ser aplicada corresponde ao primeiro elemento do vetor $\Delta\hat{\mathbf{u}}^*$, isto é:

  $$\Delta u(k) = \Delta\hat{u}^*(k|k) = K_{MPC}(\mathbf{r} - \mathbf{f})$$

  sendo K_{MPC} a primeira linha da matriz $(G^T G + \rho I)^{-1} G^T$.
- Vale notar que $\mathbf{r} = \mathbf{f} \Rightarrow \Delta u(k) = 0$, ou seja, se a resposta livre seguir o sinal de referência, não é necessário promover alterações no controle.

Interpretação para $N = M = 1$

Neste caso, tem-se $G = g(1)$, $\mathbf{r} = y_{ref}$, $\mathbf{f} = f(k+1|k)$. Logo:

$$(G^T G + \rho I)^{-1} G^T = \frac{g(1)}{g^2(1) + \rho} = K_{MPC}$$

$$\Delta u(k) = K_{MPC}(y_{ref} - f(k+1|k))$$

Observações

$$\Delta \hat{\mathbf{u}}^* = (G^T G + \rho I)^{-1} G^T (\mathbf{r} - \mathbf{f})$$

- Se a referência for constante, será empregada a seguinte notação:

$$\mathbf{r} = [y_{ref}]_N$$

sendo $[\bullet]_N$ um vetor-coluna formado pelo empilhamento vertical de N cópias de $\bullet$.

Observações

Se a referência variar com o tempo, mas for previamente conhecida, basta fazer

$$\mathbf{r} = \begin{bmatrix} y_{ref}(k+1) \\ y_{ref}(k+2) \\ \vdots \\ y_{ref}(k+N) \end{bmatrix}$$

DMC: Resumo

Informação requerida sobre a planta:

- Resposta a degrau $g(n), n = 1, 2, \ldots, N_s$
  (assume-se $g(0) = 0$ e $g(n) = g(N_s), \forall n \geq N_s$).

Parâmetros de projeto:

- Peso do controle ρ
- Horizonte de predição N
- Horizonte de controle M

Inicialização:

- Fazer $G = \begin{bmatrix} g(1) & 0 & \cdots & 0 \\ g(2) & g(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g(N) & g(N-1) & \cdots & g(N-M+1) \end{bmatrix}$
- Calcular $K_{MPC} = [\, 1\ 0\ \cdots\ 0 \,](G^T G + \rho I_M)^{-1} G^T$
- Fazer $k = 0$, $u(-1) = 0$ e $\Delta u(-1) = \Delta u(-2) = \cdots = \Delta u(-N_s + 1) = 0$

Rotina principal:

1. Ler $y(k)$ (saída da planta) e y_{ref} (valor de referência)
2. Fazer $\mathbf{r} = [y_{ref}]_N$
3. Calcular

$$f(k+i|k) = y(k) + \sum_{n=1}^{N_s-1} [g(n+i) - g(n)]\Delta u(k-n), \ i = 1, 2, \ldots, N$$

4. Fazer $\mathbf{f} = [f(k+1|k) \ f(k+2|k) \ \cdots \ f(k+N|k)]^T$
5. Calcular o incremento no controle: $\Delta u(k) = K_{MPC}(\mathbf{r} - \mathbf{f})$
6. Atualizar o controle aplicado à planta: $u(k) = u(k-1) + \Delta u(k)$
7. Fazer $k = k + 1$
8. Aguardar o próximo instante de amostragem e retornar ao passo 1

Resumo da aula de hoje

- DMC: Obtenção da equação de predição a partir da resposta a degrau da planta:

$$\hat{\mathbf{y}} = G\Delta\hat{\mathbf{u}} + \mathbf{f}$$

- Solução do problema de otimização

$$\Delta\hat{\mathbf{u}}^* = (G^T G + \rho I)^{-1} G^T (\mathbf{r} - \mathbf{f})$$

$$\Delta u(k) = \Delta\hat{u}^*(k|k) = [\, 1 \; 0 \; \cdots \; 0 \,](G^T G + \rho I)^{-1} G^T (\mathbf{r} - \mathbf{f})$$

Tópicos da próxima aula

- DMC: Sintonia de parâmetros (Período de amostragem, M, N, ρ)
- Implementação em Matlab/Simulink
- Exemplos