Funções de Matriz Quadrada

Matrizes quadradas $A \in \mathbb{R}^{n \times n}$ estão associadas a transformações lineares $f \ : \ \mathbb{R}^n \rightarrow \mathbb{R}^n$

Defina

$$A^0 = \mathbf{I} \ \ ; \ \ A^k = AA \cdots A \ (k \text{ vezes}) \quad ; \quad k \in \mathbb{Z}$$

Funções Polinomiais

Seja $f(\lambda)$ um polinômio em λ de grau finito. Por exemplo,

$$f(\lambda) = \lambda^2 + 5\lambda + 6 = (\lambda + 2)(\lambda + 3)$$

Uma função $f(A)$ é definida como

$$f(A) \triangleq A^2 + 5A + 6\mathbf{I} = (A + 2\mathbf{I})(A + 3\mathbf{I})$$

Em particular, se A assume a forma bloco diagonal

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

com A_1 e A_2 matrizes quadradas de qualquer ordem, tem-se

$$A^k = \begin{bmatrix} A_1^k & 0 \\ 0 & A_2^k \end{bmatrix} \quad ; \quad f(A) = \begin{bmatrix} f(A_1) & 0 \\ 0 & f(A_2) \end{bmatrix}$$

Como é sempre possível escrever $A = Q\hat{A}Q^{-1}$ ($\hat{A}$ é a representação de A na Forma Canônica de Jordan)

$$f(A) = f(Q\hat{A}Q^{-1}) = (Q\hat{A}Q^{-1})(Q\hat{A}Q^{-1}) + 5(Q\hat{A}Q^{-1}) + 6(QQ^{-1})$$

$$= Q\left[\hat{A}^2 + 5\hat{A} + 6\mathbf{I}\right]Q^{-1}$$

$$= Qf(\hat{A})Q^{-1}$$

ou $f(\hat{A}) = Q^{-1}f(A)Q$.

O **polinômio mínimo** de A é definido como o polinômio mônico (maior coeficiente igual a 1) $\phi(\lambda)$ de menor grau tal que $\phi(A) = \mathbf{0}$.

Portanto, $f(A) = \mathbf{0}$ se e somente se $f(\hat{A}) = \mathbf{0}$ (matrizes similares têm o mesmo polinômio mínimo).

O polinômio mínimo de uma matriz na forma de Jordan pode ser obtido por inspeção.

Se λ_i é um autovalor de A com multiplicidade n_i, o polinômio característico de A é dado por

$$\Delta(\lambda) = \det(\lambda\mathbf{I} - A) = \prod_i (\lambda - \lambda_i)^{n_i}$$

Supondo que a forma de Jordan de A é conhecida, define-se como o **índice** de λ_i a maior ordem de todos os blocos de Jordan associados a λ_i (denotado $\bar{n}_i$).

Por exemplo, λ_1 tem multiplicidade 4 nas quatro matrizes abaixo:

$$\hat{A}_1 = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix} \quad ; \quad \hat{A}_2 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}$$

$$\hat{A}_3 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix} \quad ; \quad \hat{A}_4 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}$$

$\Longrightarrow$ os índices do autovalor λ_1 são, respectivamente, 1, 2, 3 e 4.

Usando os índices $\bar{n}_i$ de todos os autovalores λ_i, o polinômio mínimo pode ser expresso da seguinte forma:

$$\phi(\lambda) = \prod_i (\lambda - \lambda_i)^{\bar{n}_i}$$

com grau $\bar{n} = \sum \bar{n}_i \leq \sum n_i = n =$ dimensão de A.

Nas matrizes acima, os polinômios mínimos são:

$$\phi_1 = (\lambda - \lambda_1) \;\; ; \;\; \phi_2 = (\lambda - \lambda_1)^2 \;\; ; \;\; \phi_3 = (\lambda - \lambda_1)^3 \;\; ; \;\; \phi_4 = (\lambda - \lambda_1)^4$$

O polinômio característico é sempre $\Delta(\lambda) = (\lambda - \lambda_1)^4$

Portanto, o polinômio mínimo é um fator do polinômio característico (se os autovalores são todos distintos, os dois polinômios são iguais).

Considere por exemplo a forma de Jordan dada por

$$\hat{A} = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}$$

Note que

$$(\hat{A} - \lambda \mathbf{I}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad ; \quad (\hat{A} - \lambda \mathbf{I})^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(\hat{A} - \lambda \mathbf{I})^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad ; \quad (\hat{A} - \lambda \mathbf{I})^k = \mathbf{0} \ \ \text{para } k \geq 4$$

Portanto, $\phi(A) = \mathbf{0}$ e nenhum outro polinômio de grau menor verifica a condição.

Teorema de Cayley-Hamilton

Seja $\Delta(\lambda) = \det(\lambda\mathbf{I} - A) = \lambda^n + \alpha_1\lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n$ o polinômio característico de A. Então,

$$\Delta(A) = A^n + \alpha_1 A^{n-1} + \cdots + \alpha_{n-1}A + \alpha_n\mathbf{I} = \mathbf{0}$$

isto é, toda matriz $A \in \mathbb{R}^{n\times n}$ satisfaz sua equação característica.

Como $n_i \geq \bar{n}_i$, o polinômio característico contém o polinômio mínimo como um fator, ou seja, para algum polinômio $h(\lambda)$

$$\Delta(\lambda) = \phi(\lambda)h(\lambda)$$

Como $\phi(A) = \mathbf{0}$,

$$\Delta(A) = \phi(A)h(A) = \mathbf{0}h(A) = \mathbf{0}$$

Pelo Teorema, A^n pode ser escrito como uma combinação linear de $\{\mathbf{I}, A, \ldots, A^{n-1}\}$.

De fato, multiplicando-se $\Delta(A) = \mathbf{0}$ por A tem-se que A^{n+1} pode ser escrito como combinação linear de $\{A, A^2 \ldots, A^n\}$, que por sua vez pode se escrever como combinação linear de $\{\mathbf{I}, A, \ldots, A^{n-1}\}$, e assim sucessivamente.

Para qualquer polinômio $f(\lambda)$, independentemente do grau, e valores apropriados de β_i, $f(A)$ pode ser expresso na forma

$$f(A) = \beta_0\mathbf{I} + \beta_1 A + \cdots + \beta_{n-1}A^{n-1}$$

Na verdade, se o polinômio mínimo (grau $\bar{n}$) de A é conhecido, A pode ser expressa como combinação linear de $\{\mathbf{I}, A, \ldots, A^{\bar{n}-1}\}$.

Expressando um polinômio qualquer $f(\lambda)$ na forma

$$f(\lambda) = q(\lambda)\Delta(\lambda) + h(\lambda)$$

$q(\lambda)$: quociente da divisão por $\Delta(\lambda)$

$h(\lambda)$: resto da divisão (grau menor que n)

$$f(A) = q(A)\Delta(A) + h(A) = q(A)\mathbf{0} + h(A) = h(A)$$

Uma alternativa à divisão de polinômios acima é dada a seguir. Defina $h(\lambda)$ como

$$h(\lambda) = \beta_0 + \beta_1\lambda + \cdots + \beta_{n-1}\lambda^{n-1}$$

β_i, $i = 0, \ldots, n-1$: incógnitas a serem obtidas

Se os n autovalores de A são distintos, β_i podem ser obtidos diretamente das n equações

$$f(\lambda_i) = q(\lambda_i)\Delta(\lambda_i) + h(\lambda_i) = h(\lambda_i) \quad , \quad i = 1, 2, \ldots, n$$

Se A tem autovalores com multiplicidade maior do que 1, a expressão acima tem que ser diferenciada (em relação a λ) para fornecer novas equações.

Seja $f(\lambda)$ uma função dada e seja $A \in \mathbb{R}^{n\times n}$ com o polinômio característico

$$\Delta(\lambda) = \prod_{i=1}^{m}(\lambda - \lambda_i)^{n_i} \quad ; \quad n = \sum_{i=1}^{m} n_i$$

Defina o polinômio de grau $n-1$ (com n coeficientes a determinar):

$$h(\lambda) = \beta_0 + \beta_1\lambda + \cdots + \beta_{n-1}\lambda^{n-1}$$

Os n coeficientes β_i podem ser obtidos do conjunto de n equações dadas por

$$f^{(\ell)}(\lambda_i) = h^{(\ell)}(\lambda_i) \qquad \left\{ \begin{array}{l} \ell = 0, 1, \ldots, n_i - 1 \\ i = 1, 2, \ldots, m \end{array} \right.$$

sendo que $f^{(\ell)}(\lambda_i) \triangleq \left.\dfrac{d^\ell f(\lambda)}{d\lambda^\ell}\right|_{\lambda=\lambda_i} \quad ; \quad h^{(\ell)}(\lambda_i) \triangleq \left.\dfrac{d^\ell h(\lambda)}{d\lambda^\ell}\right|_{\lambda=\lambda_i}$

Neste caso, $f(A) = h(A)$ e diz-se que $h(\lambda)$ é igual a $f(\lambda)$ no espectro (conjunto dos autovalores) de A.

• Dois polinômios que tenham os mesmos valores no espectro de A definem a mesma função matricial.

• O resultado acima pode ser usado para qualquer função $f(\lambda)$ (não necessariamente polinomial), definindo-se $f(A) = h(A)$ e computando-se os coeficientes β_i, $i = 0, \ldots, n-1$.

• Qualquer polinômio $h(\lambda)$ de grau $n-1$, com n parâmetros independentes, poderia ser usado.

- A^{100} , $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$; $\Delta(\lambda) = \det(\lambda \mathbf{I} - A) = (\lambda - 1)^2$

$$g(\lambda) = \alpha_0 + \alpha_1 \lambda$$

Espectro de A: $\lambda = 1$, $f(\lambda) = \lambda^{100}$

$$\left.\begin{aligned} f(1) = g(1) \;&\rightarrow\; 1^{100} = \alpha_0 + \alpha_1 \\ \frac{d}{d\lambda} f(1) = \frac{d}{d\lambda} g(1) \;&\rightarrow\; 100(1^{99}) = \alpha_1 \end{aligned}\right\} \Longrightarrow \alpha_0 = -99\,;\; \alpha_1 = 100$$

$$A^{100} = f(A) = g(A) = \alpha_0 \mathbf{I} + \alpha_1 A = \begin{bmatrix} 1 & 200 \\ 0 & 1 \end{bmatrix}$$

- Calcule $\exp(At)$, isto é, se $f(\lambda) = \exp(\lambda t)$, encontre $f(A)$

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix} \quad ; \quad \Delta(\lambda) = (\lambda - 1)^2 (\lambda - 2) \quad , \quad n_1 = 2, n_2 = 1$$

$$g(\lambda) = \alpha_0 + \alpha_1 \lambda + \alpha_2 \lambda^2 \;\; ; \;\; g(A) = \alpha_0 \mathbf{I} + \alpha_1 A + \alpha_2 A^2 = f(A)$$

$$\begin{aligned} f(1) = g(1) \quad &\rightarrow \quad \exp(t) = \alpha_0 + \alpha_1 + \alpha_2 \\ f'(1) = g'(1) \quad &\rightarrow \quad t\exp(t) = \alpha_1 + 2\alpha_2 \\ f(2) = g(2) \quad &\rightarrow \quad \exp(2t) = \alpha_0 + 2\alpha_1 + 4\alpha_2 \end{aligned}$$

$$\alpha_0 = -2t\exp(t) + \exp(2t) \quad ; \quad \alpha_1 = 3t\exp(t) + 2\exp(t) - 2\exp(2t)$$

$$\alpha_2 = \exp(2t) - \exp(t) - t\exp(t)$$

$$f(A) = \begin{bmatrix} 2\exp(t) - \exp(2t) & 0 & 2\exp(t) - 2\exp(2t) \\ 0 & \exp(t) & 0 \\ -\exp(t) + \exp(2t) & 0 & 2\exp(2t) - \exp(t) \end{bmatrix}$$

Exemplo: obtenha a expressão de $f(\hat{A})$ para

$$\hat{A} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}$$

O polinômio característico é dado por $\Delta(\lambda) = (\lambda - \lambda_1)^4$. Escolhendo (de maneira conveniente)

$$h(\lambda) = \beta_0 + \beta_1(\lambda - \lambda_1) + \beta_2(\lambda - \lambda_1)^2 + \beta_3(\lambda - \lambda_1)^3$$

A condição $f(\lambda) = h(\lambda)$ no espectro de $\hat{A}$ fornece

$$\beta_0 = f(\lambda_1) \ ; \ \ \beta_1 = \frac{f'(\lambda_1)}{1!} \ ; \ \ \beta_2 = \frac{f''(\lambda_1)}{2!} \ ; \ \ \beta_3 = \frac{f^{(3)}(\lambda_1)}{3!}$$

Assim,

$$f(\hat{A}) = f(\lambda_1)\mathbf{I} + \frac{f'(\lambda_1)}{1!}(\hat{A} - \lambda_1\mathbf{I}) + \frac{f''(\lambda_1)}{2!}(\hat{A} - \lambda_1\mathbf{I})^2 + \frac{f^{(3)}(\lambda_1)}{3!}(\hat{A} - \lambda_1\mathbf{I})^3$$

Usando as propriedades de $(\hat{A} - \lambda\mathbf{I})^k$

$$f(\hat{A}) = \begin{bmatrix} f(\lambda_1) & f'(\lambda_1)/1! & f''(\lambda_1)/2! & f^{(3)}(\lambda_1)/3! \\ 0 & f(\lambda_1) & f'(\lambda_1)/1! & f''(\lambda_1)/2! \\ 0 & 0 & f(\lambda_1) & f'(\lambda_1)/1! \\ 0 & 0 & 0 & f(\lambda_1) \end{bmatrix}$$

Por exemplo, para $f(\lambda) = \exp(\lambda t)$

$$\exp(\hat{A}t) = \begin{bmatrix} \exp(\lambda_1 t) & t\exp(\lambda_1 t) & t^2\exp(\lambda_1 t)/2! & t^3\exp(\lambda_1 t)/3! \\ 0 & \exp(\lambda_1 t) & t\exp(\lambda_1 t) & t^2\exp(\lambda_1 t)/2! \\ 0 & 0 & \exp(\lambda_1 t) & t\exp(\lambda_1 t) \\ 0 & 0 & 0 & \exp(\lambda_1 t) \end{bmatrix}$$

Exemplo: considere a matriz (com dois blocos de Jordan)

$$A = \left[\begin{array}{ccc|cc} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ \hline 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{array}\right]$$

- Se $f(\lambda) = \exp(\lambda t)$, então

$$\exp(At) = \begin{bmatrix} \exp(\lambda_1 t) & t\exp(\lambda_1 t) & t^2\exp(\lambda_1 t)/2! & 0 & 0 \\ 0 & \exp(\lambda_1 t) & t\exp(\lambda_1 t) & 0 & 0 \\ 0 & 0 & \exp(\lambda_1 t) & 0 & 0 \\ 0 & 0 & 0 & \exp(\lambda_2 t) & t\exp(\lambda_2 t) \\ 0 & 0 & 0 & 0 & \exp(\lambda_2 t) \end{bmatrix}$$

- Se $f(\lambda) = (s-\lambda)^{-1}$, então

$$(s\mathbf{I}-A)^{-1} = \begin{bmatrix} \frac{1}{(s-\lambda_1)} & \frac{1}{(s-\lambda_1)^2} & \frac{1}{(s-\lambda_1)^3} & 0 & 0 \\ 0 & \frac{1}{(s-\lambda_1)} & \frac{1}{(s-\lambda_1)^2} & 0 & 0 \\ 0 & 0 & \frac{1}{(s-\lambda_1)} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{(s-\lambda_2)} & \frac{1}{(s-\lambda_2)^2} \\ 0 & 0 & 0 & 0 & \frac{1}{(s-\lambda_2)} \end{bmatrix}$$

Teorema de Cayley-Hamilton

Seja $\Delta(\lambda) = \det(\lambda \mathbf{I} - A) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n$ o polinômio característico de A. Então,

$$\Delta(A) = A^n + \alpha_1 A^{n-1} + \cdots + \alpha_{n-1} A + \alpha_n \mathbf{I} = \mathbf{0}$$

• Se A é uma matriz diagonalizável, então existe uma transformação de similaridade dada pela matriz Q tal que

$$A = Q \Lambda Q^{-1} \quad ; \quad \Lambda \ \text{diagonal}$$

Neste caso,

$$A^2 = (Q\Lambda Q^{-1})(Q\Lambda Q^{-1}) = Q\Lambda^2 Q^{-1}$$

$$A^3 = Q\Lambda^3 Q^{-1} \ , \ \ldots \ , \ A^k = Q\Lambda^k Q^{-1}$$

Assim,

$$\Delta(A) = Q[\Lambda^n + \alpha_1 \Lambda^{n-1} + \cdots + \alpha_{n-1}\Lambda + \alpha_n \mathbf{I}]Q^{-1}$$

e cada termo dentro dos colchetes é uma matriz diagonal cujo elemento (i, i) é dado por

$$\lambda_i^n + \alpha_1 \lambda_i^{n-1} + \cdots + \alpha_{n-1}\lambda_i + \alpha_n = \Delta(\lambda_i) = 0$$

pois λ_i é um autovalor de A.

• O teorema de Cayley-Hamilton fornece uma fórmula explícita para o cálculo da matriz inversa

$$A^{-1} = -\frac{1}{\alpha_n}[A^{n-1} + \alpha_1 A^{n-2} + \cdots + \alpha_{n-1}\mathbf{I}]$$

No caso geral, qualquer matriz A sempre pode ser reduzida à forma de Jordan

$$\hat{A} = Q^{-1}AQ$$

e portanto

$$\Delta(A) = Q[\hat{A}^n + \alpha_1\hat{A}^{n-1} + \cdots + \alpha_{n-1}\hat{A} + \alpha_n\mathbf{I}]Q^{-1}$$

Para mostrar que a matriz dentro dos colchetes vale sempre zero, note que a forma de Jordan é composta de blocos diagonais $\hat{A}_i$

$$\hat{A} = \text{ diag } \{\hat{A}_1, \hat{A}_2, \ldots, \hat{A}_r\} \quad ; \quad \hat{A}^k = \text{ diag } \{\hat{A}_1^k, \hat{A}_2^k, \ldots, \hat{A}_r^k\}$$

Considerando um típico bloco $\hat{A}_i$, é preciso provar que

$$[\hat{A}_i^n + \alpha_1\hat{A}_i^{n-1} + \cdots + \alpha_{n-1}\hat{A}_i + \alpha_n\mathbf{I}] = \mathbf{0}$$

Note que os termos abaixo da diagonal principal são sempre iguais a zero, e que na diagonal principal um elemento típico (i,i) é dado por $\Delta(\lambda_i) = 0$. Na diagonal acima da diagonal principal (verificar), um termo típico é dado por

$$n\lambda_i^{n-1} + \alpha_1(n-1)\lambda_i^{n-2} + \cdots + \alpha_{n-1} = \left.\frac{d\Delta(\lambda)}{d\lambda}\right|_{\lambda=\lambda_i} = 0$$

e a raiz em questão tem multiplicidade maior do que 1. Se $\hat{A}_i$ é um bloco de Jordan de tamanho $p \times p$, λ_i necessariamente é uma raiz de no mínimo ordem p, e assim as derivadas de ordem até $p-1$ são todas iguais a zero em $\lambda = \lambda_i$. As sucessivas diagonais acima possuem termos que são múltiplos dessas derivadas da equação característica, e portanto $[\hat{A}_i^n + \alpha_1\hat{A}_i^{n-1} + \cdots + \alpha_n\mathbf{I}] = \mathbf{0}$