FUNDAÇÃO GETÚLIO VARGAS

ESCOLA DE MATEMÁTICA APLICADA

MESTRADO 2016.1

ÁLGEBRA LINEAR E APLICAÇÕES

Prof Moacyr

Resolução dos Exercícios da Lista 6

KIZZY TERRA

RIO DE JANEIRO

MAIO DE 2016

1 Verificação de Conceitos

1. Qual é a definição de uma matriz positiva definida? E semi-definida?

Uma matriz A é chamada de positiva definida se sua forma quadrática $f(x) = x^T Ax$ tem a seguinte propriedade:

- $f(x) > 0$ para todo $x \neq 0$

Uma matriz A é chamada de positiva semidefinida se sua forma quadrática $f(x) = x^T Ax$ tem a seguinte propriedade:

- $f(x) \geq 0$ para todo x

2. O que podemos dizer sobre os autovalores de uma matriz simétrica e positiva definida? E se a matriz não for simétrica?

Uma matriz simétrica e positiva definida possui autovalores reais e positivos. Se a matriz não for simétrica não podemos garantir que seus autovalores serão reais (Exemplo: Rotação de 30 graus)

3. Qual é a decomposição SVD da matriz $A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$**?**

Para calcular a decomposição SVD de uma matriz $A = USV^T$ devemos seguir os seguintes passos:

Passo 1: Calcular a matriz $A^T A$

$$\begin{pmatrix} 5 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 25 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Passo 2: Encontrar seus autovalores para construir a matriz S

Nossa matriz é diagonal e seus autovalores são os elementos da diagonal principal $\lambda_1 = 25$, $\lambda_2 = 4$ e $\lambda_3 = 0$

Sendo assim os valores singulares não nulos serão $\sigma_1 = \sqrt{25} = 5$ e $\sigma_2 = \sqrt{4} = 2$.

S é a matriz $2x3$ tal que:

$$S = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$$

Passo 3: Encontrar uma base ortonormal formada pelos autovetores associados aos autovalores de $A^T A$ para construir a matriz V^T_{3x3}

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Neste caso $V = V^T$:

$$V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Passo 4: Podemos calcular a matriz U_{2x2} de duas maneiras: calculando os autovetores da matriz AA^T ou calculando $u_i = \frac{Av_i}{\sigma_i}$.

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{5}\begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$u_2 = \frac{Av_2}{\sigma_2} = \frac{1}{2}\begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$A = USV^T$$

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

2 Exercícios I

1. Verdadeiro ou falso

(a) O conjunto das matrizes $n \times n$ positivas definidas é um cone convexo no espaço vetorial das matrizes $n \times n$

Falso.

Definição: Um cone convexo é um conjunto C tal que, para todo $x_1, x_2 \in C$ e $\alpha_1, \alpha_2 \geq 0$ $\alpha_1 x_1 + \alpha_2 x_2 \in C$

Para verificar isto devemos verificar se o espaço das matrizes positivas definidas é fechado para soma e para produto por escalar não negativo:

i) Fechado para soma:

Sejam A e B matrizes positivas definidas tais que $x^T Ax > 0$ e $x^T Bx > 0$ então:

$$x^T Ax + x^T Bx > 0 \Rightarrow x^T (A + B)x > 0$$

i) Fechado para o produto por escalar não negativo:

Seja A uma matriz positiva definida tal que $x^T Ax > 0$ então:

$$x^T \alpha Ax > 0 \Leftrightarrow \alpha > 0$$

Portanto o espaço não é fechado para o produto por escalar real, pois a propriedade não se preserva para $\alpha = 0$.

(b) O conjunto das matrizes $n \times n$ positivas definidas é um subespaço do espaço das matrizes $n \times n$

Falso.

Verificando se o espaço das matrizes positivas definidas é fechado para o produto por escalar:

Seja A uma matriz positiva definida então $x^T Ax > 0$ então:

Seja $\alpha \in \mathbb{R}$ então $x^T(\alpha A)x > 0$ se e somente se $\alpha > 0$ portanto o espaço não é fechado para o produto por escalar real, pois a propriedade não se preserva para escalares não positivos.

(c) Os elementos da diagonal principal de uma matriz positiva definida são números positivos

Verdadeiro.

Tome o produto da matriz A positiva definida de ordem n pelos n vetores e_i canônicos:

$$e_i^T A e_i = a_{ii} > 0$$

Teremos que todos os elementos da diagonal principal da matriz A deverá ser positivo.

(d) Se uma matriz quadrada é diagonal dominante então ela é positiva definida

Falso.

Contra-exemplo:

$$A = \begin{bmatrix} -4 & 2 \\ 1 & -2 \end{bmatrix}$$

Esta matriz é diagonal dominante mas não é positiva definida.

(e) Existe uma matriz positiva definida 2×2 com dois elementos negativos e dois positivos

Verdadeiro.

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

$$p_A(\lambda) = (2-\lambda)(2-\lambda) - 1 = (3 - 4\lambda + \lambda^2)$$

$$\lambda_1 = 3 \text{ e } \lambda_2 = 1$$

A matriz é positiva definida e possui dois números negativos.

(f) Se os determinantes dos menores principais de uma matriz M são todos positivos então $M > 0$. Se alternam o sinal então $M < 0$.

Falso.

A primeira afirmação é verdadeira, entretanto para $M < 0$ os menores principais devem alternar o sinal, mas além disso é necessário que o primeiro menor principal seja negativo.

2. Seja $A_{n\times n}$ uma matriz diagonalizável e$p_A(\lambda) = det(A - \lambda I)$ seu polinômio característico. Mostre que se no polinômio $p_A(\lambda)$ substituirmos a variável λ por A então $p_A(A) = 0_{n\times n}$, isto é, obteremos a matriz nula (teorema de Cayley-Hamilton). O que podemos dizer sobre matrizes que não são diagonalizáveis?

$$p_A(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_2\lambda^2 + a_1\lambda + a_0$$

$$p_A(A) = a_n A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0$$

Se A é uma matriz diagonalizável tal que $A = PDP^{-1}$ então:

$$p_A(A) = a_n PD^nP^{-1} + a_{n-1}PA^{n-1}P^{-1} + \cdots + a_2PA^2P^{-1} + a_1PAP^{-1} + a_0PP^{-1}$$

$$p_A(A) = P.(a_nD^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A + a_0).P^{-1}$$

$$p_A(A) = P.\begin{pmatrix} a_n\lambda_1^n + \cdots + a_2\lambda_1^2 + a_1\lambda_1 + a_0 & 0 & 0 & \cdots \\ 0 & a_n\lambda_2^n + \cdots + a_2\lambda_2^2 + a_1\lambda_2 + a_0 & 0 & \cdots \\ 0 & 0 & a_n\lambda_3^n + \cdots + a_2\lambda_3^2 + a_1\lambda_3 + a_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \\ 0 & 0 & 0 & \cdots \end{pmatrix}$$

$$p_A(A) = P.\begin{pmatrix} p_A(\lambda_1) & 0 & 0 & \cdots & 0 \\ 0 & p_A(\lambda_2) & 0 & \cdots & 0 \\ 0 & 0 & p_A(\lambda_3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & p_A(\lambda_n) \end{pmatrix}.P^{-1}$$

$$p_A(A) = P.\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.P^{-1} = 0_{nxn}$$

3. Seja$F : R^n \to R$ uma função que tem todas as segundas derivadas contínuas ($F \in C^2$). Mostre que se $x^* \in \mathbb{R}^n$é um ponto onde a derivada se anula e a hessiana de F neste ponto tem todos os autovalores positivos então x^* é um ponto de mínimo local de F.

A matriz hessiana F é uma matriz simétrica e dado que ela tem os autoalores podsitivos podemos garantir que ela é uma matriz positiva definida. Para mostrar que x^* é ponto mínimo devemos analisar a expansão de Taylor de ordem 2.

4. Mostre que se a matriz A é simétrica e v é um autovetor de A então o complemento ortogonal do espaço gerado por v é invariante por A.

Definição: W, subespaço vetorial de V, é dito invariante sob o operador $T : V \to V$ se $T(W) \subset W$.

Queremos mostrar que para todo u pertencente ao complemento ortogonal do espaço gerado por v Au também pertence ao complemento ortogonal do espaço gerado por v.

Temos que

$$Av = \lambda v$$

Para todo u pertencente ao complemento ortogonal do espaço gerado por v:

$$u^T v = v^T u = 0$$

Se A é simétrica:

$$u^T Av = u^T A^T v = (Au)^T v = v^T (Au)$$

$$u^T Av = u^T (\lambda v) = \lambda u^T v = 0$$

$$u^T Av = v^T (Au) = 0$$

Logo $(Au) \in ger(v)^{\perp}$

5. A decomposição em valores singulares (forma reduzida) da matriz da transformação $A : \mathbb{R}^4 \to \mathbb{R}^5$ é $U = (u_1|u_2|u_3|u_4)$, $S = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ e $V = (v_1|v_2|v_3|v_4)$, onde as colunas u_i são os vetores de $\mathbb{R}^5$ e as colunas v_i são os vetores de $\mathbb{R}^4$.

a) Considere a função $f : \mathbb{R}^4 \to \mathbb{R}$ dada por $f(w) = \frac{\|Aw\|}{\|w\|}$ exiba um vetor $w*$ que maximiza $f(w)$.

Conforme discutido na monitoria:

$$max \frac{\|Aw\|}{\|w\|} \Rightarrow max \|A \frac{w}{\|w\|}\| \Rightarrow max_{\|w\|=1} \|Aw\|$$

$$max_{\|w\|=1} \|Aw\| = \sigma_1 \Rightarrow w^* = v_1$$

σ_1, maior valor singular

v_1, autovetor associado ao maior valor singular

Para mostrar isso fazemos:

Considere $A = USV^T$ e $w = a_1 v_1 + ... + a_n v_n$:

$$\|w\|^2 = 1 \Rightarrow a_1^2 + ... + a_n^2 = 1 \ (\|w\| = 1, \|v_i\| = 1 \ ):$$

$$Aw = USV^T = U \begin{pmatrix} \sigma_1 a_1 \\ \sigma_2 a_2 \\ \vdots \\ \sigma_n a_n \end{pmatrix} \Rightarrow Aw = \sigma_1 a_1 u_1 + \cdots + \sigma_n a_n u_n$$

$$Aw = \sigma_1 a_1 u_1 + \cdots + \sigma_n a_n u_n \Rightarrow \|Aw\|^2 = \sigma_1^2 a_1^2 + \cdots + \sigma_n^2 a_n^2$$

$$\frac{\|Aw\|^2}{\|w\|^2} = \sigma_1^2 a_1^2 + \cdots + \sigma_n^2 a_n^2$$

Se σ_1 é o maior valor singular o valor w^* que irá maximizar $\frac{\|Aw\|}{\|w\|}$ é $w^* = a_1 v_1 = 1 * v_1 = v_1$

b) Determine bases ortogonais para $N(A)$, $Im(A)$, $N(A^T)$ e $Im(A^T)$

$A = USV^T$

i) $Im(A)$

$(\Rightarrow) C(A) \subseteq ger(u_1, u_2, \cdots, u_r)$ onde $u_1, u_2, \cdots, u_r$ são os autovetores de AA^T associados aos autovalores não nulos.

$(\Leftarrow) ger(u_1, u_2, \cdots, u_r) \subseteq C(U) \subseteq C(A)$

Logo $C(A) = ger(u_1, u_2, \cdots, u_r)$

ii) $N(A^T) = C(A)^\perp = ger(u_{r+1}, u_{r+2}, \cdots, u_m)$

iii) $Im(A^T)$

$(\Rightarrow) C(A^T) \subseteq ger(v_1, v_2, \cdots, v_r)$ onde $v_1, v_2, \cdots, v_r$ são os autovetores de A^TA associados aos autovalores não nulos.

$(\Leftarrow) ger(v_1, v_2, \cdots, v_r) \subseteq C(V) \subseteq C(A^T)$

Logo $C(A^T) = ger(v_1, v_2, \cdots, v_r)$

iv) $N(A) = C(A^T)^\perp = ger(v_{r+1}, v_{r+2}, \cdots, v_n)$

6. Calcule a decomposição SVD da matriz $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix}$. Use esta decomposição para calcular a pseudo-inversa de A.

$$A = USV^T$$

Passo 1: Calcular a matriz A^TA

$$\begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$$

Passo 2: Encontrar seus autovalores para construir a matriz S

$p_\lambda(A^TA) = (2-\lambda)(3-\lambda) = 0 \Rightarrow \lambda_1 = 3 \ e \ \lambda_2 = 2$

$$S = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}$$

Passo 3: Encontrar uma base ortonormal formada pelos autovetores associados aos autovalores de A^TA para construir a matriz V_{2x2}^T

$$v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ e } v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$V = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ 1 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Passo 4: Podemos calcular a matriz U_{3x3} de duas maneiras: calculando os autovetores da matriz AA^T ou calculando $u_i = \frac{Av_i}{\sigma_i}$

$$u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$u_2 = \frac{Av_2}{\sigma_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{1}{2} \\ -1 \end{pmatrix}$$

$$u_2 = \frac{u_2}{\|u_2\|} = \frac{2}{\sqrt{5}} \begin{pmatrix} 0 \\ -\frac{1}{2} \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix}$$

Deve-se ainda calcular um terceiro vetor u_3 que componha uma base ortnormal com u_1 e u_2. Para fazer isso podemos utilizar Gram-Schimdt.

$$A = USV^T$$

7. Considere os pontos $P_1 = (x_1, y_1), ..., P_n = (x_n, y_n)$. Queremos encontrar uma reta e um erro $y = ax + b + \varepsilon$ que melhor se ajustam aos pontos dados.

a) Mostre como podemos usar a decomposição em valores para encontrar os parâmetros a e b que minimizam $\|\varepsilon\|_2$

Referência: Páginas 553 e 554 do Poole.