MATD41-Introdução aos Modelos Lineares

1.4 – Formas Quadráticas

Kim Samejima

IME-UFBA

Formas quadráticas

Seja $\mathbf{x}$ vetor coluna n-dimensional e A_n matriz quadrada de dimensão n.
Uma forma quadrática será uma expressão da forma

$$\mathbf{x}^\top A\mathbf{x} = \sum_{i=1}^{n}\sum_{i=1}^{n} a_{ij}x_i x_j$$

Casos particulares:

$$\mathbf{x}^\top I\mathbf{x} = \mathbf{x}^\top\mathbf{x} = \sum_{i=1}^{n} x_i^2,$$

$$\mathbf{x}^\top J\mathbf{x} = \left(\sum_{i=1}^{n} x_i\right)^2,$$

Se A é simétrica:

$$\mathbf{x}^\top A\mathbf{x} = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_i x_j$$

Matrizes positivas definidas

Seja A_n matriz quadrada. A será chamada *positiva semidefinida* ou *não negativa definida* SSE:

- A é simétrica;
- para qualquer vetor real $\mathbf{x}$, vale que

$$\mathbf{x}^\top A\mathbf{x} \geq 0$$

Se $\mathbf{x}^\top A\mathbf{x} = 0$ somente se $\mathbf{x} = 0$ e $\mathbf{x}^\top A\mathbf{x} > 0$ para qualquer $\mathbf{x} \neq 0$, então A é chamada *positiva definida*

Exercício

Mostre que se uma matriz quadrada e simétrica de ordem n pode ser escrita como $\Sigma = L^\top L$ para algum $L \neq 0$, então Σ possui todos os seus autovalores maiores que zero.

Autovalores:

Todos os autovalores de uma matriz não-negativa definida são não negativos.

Teoremas

Teorema 1

Se A e B são matrizes positivas definidas e se B-A é positiva semi-definida, então $A^{-1} - B^{-1}$ *é positiva semi definida.*

Teorema 2

As seguintes afirmações são equivalentes:

- $\mathbf{x}^\top A\mathbf{x} > 0$ *para todo* $\mathbf{x} \neq 0$*;*
- *Todos os autovalores de A são positivos;*
- *Os menores principais de A são todos positivos;*
- $A = L^\top L$ *para alguma L não singular(Decomposição de Cholesky).*

Distância de Mahalanobis I

- Distância entre $\mathbf{x}$ e μ, mas se $cov(\mathbf{X}) = \Sigma = diag(0.5, 2)$ e considerarmos a distância euclidiana, então $(x_1 - \mu_1)^2$ terá, em geral, resultados mais condensados que $(x_2 - \mu_2)^2$. Logo, a distância euclidiana só faz sentido se $\Sigma = I$.

- É preciso então ponderar as distâncias pelo seu peso (inverso) na matriz de covariâncias:

$$d_\Sigma(\mathbf{x}, \mu) = \Big((\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu)\Big)^{1/2}$$

  que é chamada *distância de Mahalanobis* entre $\mathbf{x}$ e μ.
- Se $\Sigma = I$ então $d_\Sigma(\mathbf{x}, \mu)$ coincide com a distância euclidiana.
- Exemplo no Itempool.

Distância de Mahalanobis II

Exemplos:

$\Sigma = diag(0.5, 2)$

$$\Rightarrow \quad d_{\Sigma}(\mathbf{x}, \mu) = \left(2(x_1 - \mu)^2 + 0.5(x_2 - \mu)^2\right)^{1/2},$$

$$\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

$$\Rightarrow \quad d_{\Sigma}(\mathbf{x}, \mu) = \left(4/3((x_1 - \mu_1)^2 + (x_2 - \mu_2)^2 - \underbrace{(x_1 - \mu_1)(x_2 - \mu_2)}\right)^{1/2}.$$

O termo $(x_1 - \mu_1)(x_2 - \mu_2)$ regula a distância entre $\mathbf{x}$ e μ de acordo com a correlação entre X_1 e X_2.

Modelo de Regressão I

Considere o modelo de regressão linear simples abaixo:

$$y = \beta_0 + \beta_1 X + \epsilon$$

Se tivermos N observações $(x_i, y_i), i = 1, ..., N$ então podemos escrever:

$$X = \begin{bmatrix} \mathbf{1} & \mathbf{x} \end{bmatrix} \quad , \quad \mathbf{y}^\top = [y_1, ..., y_N] \quad \text{e} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

Buscamos um vetor β que minimize:

$$(\mathbf{y} - \widehat{\mathbf{y}})^\top (\mathbf{y} - \widehat{\mathbf{y}}) = (\mathbf{y} - \beta X)^\top (\mathbf{y} - \beta X)$$

Equações Normais

Se desejamos $\epsilon = (\mathbf{y} - \widehat{\mathbf{y}})$ mínimo, então $\widehat{\mathbf{y}}$ deve ser projeção ortogonal de $\mathbf{y}$.
Logo, $(\mathbf{y} - \widehat{\mathbf{y}})$ também deve ser ortogonal a $\mathbf{1}$ e a $\mathbf{x}$. Assim:

$$(\mathbf{y} - \widehat{\mathbf{y}})^\top \mathbf{1} = 0 \quad (1)$$

$$(\mathbf{y} - \widehat{\mathbf{y}})^\top \mathbf{x} = 0 \quad (2)$$

(1) e (2) são chamadas *equações normais*.