Matrizes Positivas Definidas

Álgebra Linear (MAT-27)

Ronaldo Rodrigues Pelá

IEFF-ITA

7 de novembro de 2011

Roteiro

Roteiro

Por que saber se uma matriz é definida positiva?

- Importância do sinal dos autovalores.
  - Os autovalores devem ser reais.
  - Problemas de análise estabilidade: $e^{\lambda t}$
  - A parte real de λ deve ser negativa para que a função decresça.
- Conceitos que só têm sentido para matrizes simétricas (ou hermitianas).
- Questão que surge em problemas de otimização.
  - Muito frequente em Engenharia.
  - Teste da segunda derivada.

Por que saber se uma matriz é definida positiva?

- Critério de estabilidade de Lyapunov. Dado um sistema

$$\dot{X} = f(X, t)$$

- O ponto de equilíbrio $\mathbb{O}$ é uniformemente assintoticamente estável se existir uma função potencial $V : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ com $\partial V/\partial X$ e $\partial V/\partial t$ contínuas e $V(\mathbb{O}, t) = 0$ satisfazendo a:
    1. $\alpha(||X||) \leq V(X, t) \leq \beta(||X||) \quad \forall t$, com $\alpha(.)$ e $\beta(.)$ funções não decrescentes.
    2. $\dot{V}(X, t) \leq -\gamma(||X||) \quad \forall t$, com $\gamma(.)$ função não decrescente.
    3. $V(X, t)$ é radialmente ilimitada: $V(X, t) \to \infty \quad ||X|| \to \infty$.
- Num cenário "típico": $V(X) = X^\dagger P X$ com P hermitiana.
- P (pelo menos) deve ser positiva definida.

Roteiro

Análise de um caso

- Considere as funções

$$F(x,y) = 7+2(x+y)^2-y\sin y-x^3 \qquad f(x,y) = 2x^2+4xy+y^2.$$

- Será que elas têm um mínimo em $x = y = 0$?
- Os termos de ordem zero $F(0,0) = 7$ e $f(0,0) = 0$ não ajudam a responder. Eles simplesmente representam um "shift" nos gráficos de F e f.

Análise de um caso

- Os termos lineares fornecem uma condição necessária: para que se tenha alguma chance de mínimo, as primeiras derivadas devem se anular.

$$\frac{\partial F}{\partial x} = 4(x+y) - 3x^2 \qquad \frac{\partial F}{\partial y} = 4(x+y) - y\cos y - \sin y.$$

$$\frac{\partial f}{\partial x} = 4x + 4y \qquad \frac{\partial f}{\partial y} = 4x + 2y.$$

- Isto de fato acontece!
- Portanto, $(0,0)$ é um ponto estacionário de F e de f.

Análise de um caso

- As segundas derivadas em $(0,0)$ são decisivas

$$\frac{\partial^2 F}{\partial x^2} = 4 - 6x \qquad \frac{\partial^2 f}{\partial x^2} = 4.$$

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x} = 4 \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 4.$$

$$\frac{\partial^2 F}{\partial y^2} = 4 + y \sin y - 2 \cos y \qquad \frac{\partial^2 f}{\partial x^2} = 2.$$

Análise de um caso

- Calculando em $(0,0)$

$$\frac{\partial^2 F}{\partial x^2} = \frac{\partial^2 f}{\partial x^2} = 4.$$

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 4.$$

$$\frac{\partial^2 F}{\partial y^2} = \frac{\partial^2 f}{\partial x^2} = 2.$$

- As duas funções se comportam exatamente da mesma forma quando estão próximas da origem.
- F possui um mínimo em $(0,0)$ se e somente se f possuir um mínimo em $(0,0)$.

Análise de um caso

- Os termos de mais alto grau de F não têm influência sobre $(0,0)$ ser um mínimo local.
- Mas podem influenciar no caso de a pergunta ser sobre um mínimo global.

Forma Quadrática

Toda forma quadrática $ax^2 + 2bxy + cy^2$ possui um ponto estacionário na origem.

Análise de um caso

- Se o ponto estacionário de uma $F(x, y)$ ocorrer em (α, β) e não na origem, basta efetuar a translação

$$\bar{x} = x - \alpha \qquad \bar{y} = y - \beta$$

e analisar o que ocorre com:

$$f(\bar{x}, \bar{y}) = \frac{\bar{x}^2}{2} \frac{\partial^2 F}{\partial x^2}(\alpha, \beta) + \bar{x}\bar{y} \frac{\partial^2 F}{\partial x \partial y}(\alpha, \beta) + \frac{\bar{y}^2}{2} \frac{\partial^2 F}{\partial y^2}(\alpha, \beta)$$

que representa o comportamento de $F(x, y)$ nas vizinhanças de (α, β).

Análise de um caso

- Em princípio, basta analisar o seguinte polinômio

$$ax^2 + 2bxy + cy^2$$

  obtido a partir das segundas derivadas.
- Queremos investigar se:

$$ax^2 + 2bxy + cy^2 > 0 \quad \forall (x,y) \neq (0,0) \quad \text{ponto de mínimo}$$

  ou

$$ax^2 + 2bxy + cy^2 < 0 \quad \forall (x,y) \neq (0,0) \quad \text{ponto de máximo}$$

- Pode ser que nenhuma das duas condições anteriores se verifique (ponto de sela).
- As derivadas superiores podem ser necessárias quando a parte quadrática é singular.

Análise de um caso

- Que condições garantem que

$$ax^2 + 2bxy + cy^2 > 0 \quad \forall (x, y) \neq (0, 0)$$

isto é, que seja definida positiva?

1. Se $ax^2 + 2bxy + cy^2$ é definida positiva, então $a > 0$.
2. Se $ax^2 + 2bxy + cy^2$ é definida positiva, então $c > 0$.

- As duas condições anteriores são suficientes para garantir que $ax^2 + bxy + cy^2$ seja positiva definida?
- Não (necessariamente)! Tome $x^2 - 10xy + y^2$ e aplique em $(1, 1)$: o resultado é -8.
- Então o sinal do termo misto deve ser positivo?
- Não (necessariamente)! Tome $x^2 + 10xy + y^2$ e aplique em $(-1, 1)$: o resultado é -8.

Análise de um caso

- Que condições garantem que

$$ax^2 + 2bxy + cy^2 > 0 \quad \forall (x,y) \neq (0,0)$$

isto é, que seja definida positiva?

- A condição que falta deve envolver a, b e c. Vejamos:

$$ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2$$

- Portanto, basta exigir que:

$$a > 0 \qquad \text{e} \qquad ac > b^2$$

- Note que a condição $c > 0$ é consequência das duas condições anteriores (e pode ser suprimida).

Análise de um caso

- Analogamente, as condições que garantem que

$$ax^2 + 2bxy + cy^2 < 0 \quad \forall (x, y) \neq (0, 0),$$

isto é, que seja definida negativa, são:

$$a < 0 \qquad \text{e} \qquad ac > b^2.$$

- Esta verificação fica como exercício.
- Quando $ac < b^2$, teremos um ponto de sela. Veja que é este o caso das funções $F(x, y)$ e $f(x, y) = 2x^2 + 4xy + y^2$.

Análise de um caso

- Matricialmente, teríamos:

$$ax^2 + 2bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- Definindo $u = [x \quad y]^T$ e A a matriz 2×2 acima:

$$ax^2 + 2bxy + cy^2 = u^T A u.$$

- Façamos uma mudança de coordenadas $\tilde{u} = Mu$. Logo:

$$u^T A u = \tilde{u}^T (M^T A M) \tilde{u}.$$

Análise de um caso

- Como A é simétrica, podemos tomar M formada pelos seus autovetores ortonormais, de modo que $M^T AM = D$ é uma matriz diagonal, digamos $\mathrm{diag}(\lambda_1, \lambda_2)$.
- Sendo $\tilde{u} = [\tilde{x} \quad \tilde{y}]^T$, temos

$$ax^2 + 2bxy + cy^2 = \lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2$$

- As condições que chegamos antes poderiam ser inferidas a partir dos autovalores. Veja que:
  - $\lambda_1 > 0, \lambda_2 > 0$: positiva definida.
  - $\lambda_1 < 0, \lambda_2 < 0$: negativa definida.
  - Sinais contrários: indefinida (caso equivalente ao ponto de sela).

Dimensões maiores

- Já estudamos o caso bidimensional. Vejamos agora o que acontece para dimensões maiores.
- Se tivermos uma função $F : \mathbb{R}^n \to \mathbb{R}$, o seu polinômio de Taylor em torno de $\mathbb{O}$ é:

$$F(x) = F(\mathbb{O}) + x^T \frac{\partial F}{\partial x}(\mathbb{O}) + \frac{1}{2} x^T A x + \text{termos de ordem superior}$$

- OBS.: Notação compacta para o gradiente: $\nabla F = \dfrac{\partial F}{\partial x}$.
- A matriz A é formada pelas derivadas segundas: $a_{ij} = \dfrac{\partial^2 F}{\partial x_i \partial x_j}(\mathbb{O})$. Esta é uma matriz que é simétrica (a menos de funções "teimosamente mal-comportadas"!).

Dimensões maiores

- Suponha que a origem $\mathbb{O}$ é um ponto estacionário:

$$\frac{\partial F}{\partial x}(\mathbb{O}) = \mathbb{O}$$

- Para saber se é máximo ou mínimo (ou ponto de sela), devemos analisar o sinal de $x^T Ax$, já que

$$F(x) = F(\mathbb{O}) + \frac{1}{2}x^T Ax + \text{termos de ordem superior}$$

- E se o ponto estacionário não for a origem e sim um ponto α? Neste caso, basta fazer uma tranlação, análoga a que antes fizemos: $\bar{x} = x - \alpha$.

Dimensões maiores

- Veja que

$$x^T A x = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- que resulta na seguinte forma quadrática:

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

- no nosso caso, é melhor trabalhar com a forma matricial...

Dimensões maiores

- Precisamos responder à pergunta: que matrizes simétricas têm a propriedade $x^T Ax > 0 \quad \forall x \neq \mathbb{O}$?
- Quando uma matriz satifaz a desigualdade anterior ela é dita ser positiva definida.
- Para um caso mais geral (complexo), precisamos de uma ligeira adaptação:

Matriz Positiva Definida

Uma matriz $A \in M_n(\mathbb{C})$ hermitiana é positiva definida se $x^\dagger Ax > 0 \quad \forall x \neq \mathbb{O}$.

- O caso real é análogo, bastando trocar hermitiana por simétrica e $x^\dagger$ por x^T.
- Analogamente, podemos definir o conceito de matriz positiva semidefinida, quando $x^\dagger Ax \geq 0 \quad \forall x \neq \mathbb{O}$

Dimensões maiores

- Analogamente, podemos definir o conceito de matriz negativa definida.

Matriz Negativa Definida

Uma matriz $A \in M_n(\mathbb{C})$ hermitiana é negativa definida se $x^\dagger A x < 0 \quad \forall x \neq \mathbb{O}$.

- Analogamente, podemos definir o conceito de matriz negativa semidefinida, quando $x^\dagger A x \leq 0 \quad \forall x \neq \mathbb{O}$
- Há algumas matrizes hermitianas que não são nem positivas (semi)definidas, nem negativas (semi)definidas. Este é o caso da matriz $\mathrm{diag}(1, -1)$.

Dimensões maiores

- Voltando ao problema, precisamos verificar que matrizes simétricas têm a propriedade $x^T A x > 0 \quad \forall x \neq \mathbb{O}$.
- Há diversas formas de fazer isto.
- Veremos a dedução de uma delas.
- Nesta dedução, a ideia central é fazer uma mudança de base: $\tilde{x} = Mx$, para alguma matriz M escolhida convenientemente.

$$x^T A x = \tilde{x}^T (M^T A M) \tilde{x}$$

Dimensões maiores

- Como A é simétrica (ou então hermitiana), sempre conseguimos encontrar uma base ortonormal de autovetores.
- Tomando M a matriz formada por estes autovetores, então $M^T AM$ será uma matriz diagonal.

$$x^T Ax = \lambda_1 \tilde{x}_1^2 + \lambda_2 \tilde{x}_2^2 + \ldots + \lambda_n \tilde{x}_n^2$$

- Se todos os autovalores forem positivos, teremos uma matriz positiva definida.
- Se todos os autovalores forem negativos, teremos uma matriz negativa definida.

Teste de positividade

- O teste dos autovalores não é o único para verificar se uma matriz é positiva definida.
- Vimos outro teste para o caso bidimensional.
- Vejamos o resultado geral

Teste de positividade

Para uma matriz $A \in M_n(\mathbb{C})$ hermitiana as seguintes afirmações são equivalentes

- $x^\dagger A x > 0 \quad \forall x \neq \mathbb{O}$.
- Todos os autovalores de A são positivos.
- Os menores principais de A são todos positivos.
- $A = S^2$ para alguma matriz hermitiana S não singular.
- $A = T^\dagger T$ para alguma matriz T não singular.

Teste de positividade

- OBS.: os menores principais de uma matriz A $n \times n$ são os determinantes das submatrizes

$$[a_{11}], \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{23} & a_{32} & a_{33} \end{bmatrix}, \ldots, A$$

Produto Interno

- Uma matriz hermitiana positiva definida induz um produto interno:

$$\langle x|y\rangle = x^\dagger A y$$

- De fato:
  1. Simetria: $\langle x|y\rangle^* = y^\dagger A x = \langle x|y\rangle$.
  2. Linearidade: $\langle x|y+\alpha z\rangle = x^\dagger A(y+\alpha z) = \langle x|y\rangle + \alpha\langle x|z\rangle$.
  3. Positividade: $\langle x|x\rangle = x^\dagger A x > 0$, pois A é positiva definida.

Comentário

- Pode existir uma matriz positiva definida não simétrica, ou não hermitiana?
- Esta questão não tem muito sentido...
- Uma matriz não simétrica (ou não hermitiana) pode não ser diagonalizável.

Comentário

- Uma questão mais razoável para matrizes não hermitianas A é verificar se sua parte hermitiana é positiva definida.

$$A = \frac{1}{2}(A + A^{\dagger}) + \frac{1}{2}(A - A^{\dagger})$$

- A componente $\frac{1}{2}(A + A^{\dagger})$ da soma acima é conhecida como a parte hermitiana de A.
- Testar se $\frac{1}{2}(A + A^{\dagger})$ é positiva definida garante que todos os autovalores de A tenham parte real positiva.

Roteiro

Exemplos

- A matriz
$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$$
é positiva definida
- Menores principais de A: 2, 3 e 4.

Exemplos

- As matrizes

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 2 \end{bmatrix}$$

  não são positivas definidas.
- A tem um menor principal: 0.
- B não é inversível.

Exemplos

- Há uma solução real para

$$-x^2 - 5y^2 - 9z^2 - 4xy - 6xz - 8yz = 1?$$

- Veja que

$$x^2+5y^2+9z^2+4xy+6xz+8yz = [x \quad y \quad z]\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 4 \\ 3 & 4 & 8 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = u^T A u$$

- A matriz acima tem determinante negativo.
- Ao menos um autovalor é negativo.
- Existe um certo v tal que $Av = -\alpha v$ com $\alpha > 0$. Assim, tomando $u = \beta v$, temos

$$u^T A u = -\alpha\beta^2 ||v||^2$$

- Basta tomar $\beta = 1/(||v||\sqrt{\alpha})$ que $u^T A u = -1$.

Exemplos

- A seguinte função é um produto interno no $\mathbb{R}^2$

$$\langle (x,y)|(u,v)\rangle = xu + 4yv - xv - yu?$$

- Escrevendo matricialmente

$$xu + 4yv - xv - yu = [x \quad y]\begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}$$

- A matriz é simétrica e positiva definida (pois tem menores principais: 1 e 3).
- Logo se trata sim de um produto interno.