Notações e revisão de álgebra linear

Marina Andretta

ICMC-USP

17 de agosto de 2016

Baseado no livro Introduction to Linear Optimization, de D. Bertsimas e J. N. Tsitsiklis.

Denotaremos a matriz transposta de uma matriz A por A^T.

Quando nos referirmos a vetores, a menos que seja especificado o contrário, estaremos nos referindo a um vetor coluna.

Usaremos e_i para denotar o vetor que tem 1 na componente i e 0 nas demais.

Vetores e matrizes

Se x e y são vetores em $\boldsymbol{R}^n$,

$$x^T y = y^T x = \sum_{i=1}^{n} x_i y_i$$

é chamado de produto interno.

Dois vetores são ditos ortogonais se o produto interno entre eles for 0.

Para qualquer vetor x, vale que $x^T x \geq 0$.

A igualdade vale se, e somente se, $x = 0$.

A expressão $\sqrt{x^T x}$ é a norma Euclidiana de x, denotada por $\|x\|$.

A desigualdade de Cauchy-Schwartz diz que, para quaisquer dois vetores x e y de mesma dimensão,

$$|x^T y| \leq \|x\| \|y\|,$$

com a igualdade valendo se, e somente se, um dos vetores for um múltiplo escalar do outro (ou seja, $x = \alpha y$ para algum escalar α).

Cada elemento c_{ij} do produto C de duas matrizes $A \in \boldsymbol{R}^{n \times p}$ e $B \in \boldsymbol{R}^{p \times m}$, $C = AB$, é dado por

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = a_i^T B_j,$$

com a_i a i-ésima linha de A e B_j a j-ésima coluna de B.

O produto de matrizes é associativo, ou seja, $(AB)C = A(BC)$.

Mas, no geral, não é comutativo, ou seja, geralmente $AB \neq BA$.

Vale a propriedade $(AB)^T = B^T A^T$.

Sejam $A \in \boldsymbol{R}^{m \times n}$ e $x \in \boldsymbol{R}^n$.

A i-ésima coluna de A, A_i, pode ser escrita como $Ae_i = A_i$.

O vetor x pode ser escrito como $x = \sum_{i=1}^{n} x_i e_i$.

Assim,

$$Ax = A \sum_{i=1}^{n} e_i x_i = \sum_{i=1}^{n} A e_i x_i = \sum_{i=1}^{n} A_i x_i.$$

Outra maneira de escrever o produto Ax é

$$Ax = \begin{pmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_m^T x \end{pmatrix},$$

com a_j a j-ésima linha de A.

A notação $x \geq 0$ significa que $x_i \geq 0$ para toda componente i de x.

Analogamente, $x > 0$ significa que $x_i > 0$ para toda componente i de x.

Para uma matriz A, as notações $\geq$ e $>$ têm o mesmo significado.

Inversão de matrizes

Seja A uma matriz quadrada.

Se existe uma matriz B de mesma dimensão tal que $AB = BA = I$, dizemos que A é inversível ou não-singular.

Essa matriz B é chamada de inversa de A, denotada por A^{-1}, e é única.

Vale que $(A^T)^{-1} = (A^{-1})^T$.

Se A e B são matrizes inversíveis de mesma dimensão, temos que $(AB)^{-1} = B^{-1}A^{-1}$.

Dada uma coleção de vetores $x^1, ..., x^k \in \boldsymbol{R}^n$, dizemos que eles são linearmente dependentes se existem escalares $\alpha_1, ..., \alpha_k \in \boldsymbol{R}$, não todos nulos, tal que $\sum_{i=1}^{k} \alpha_i x^i = 0$.

Caso contrário, os vetores são ditos linearmente independentes.

Uma definição equivalente para vetores linearmente independentes é que nenhum dos vetores x^i sejam uma combinação linear dos demais vetores.

Inversão de matrizes

Teorema 1: *Seja A uma matriz quadrada. Então, todas as seguintes afirmações são equivalentes:*

(a) *A é inversível.*

(b) *A^T é inversível.*

(c) *O determinante de A é diferente de zero.*

(d) *As linhas de A são linearmente independentes.*

(e) *As colunas de A são linearmente independentes.*

(f) *Para todo vetor b, o sistema linear $Ax = b$ tem solução única.*

(g) *Existe um vetor b para o qual o sistema linear $Ax = b$ tem solução única.*

Supondo que a matriz A seja inversível, a solução do sistema $Ax = b$ é dada explicitamente por $x = A^{-1}b$.

Subespaços e bases

Um subconjunto não-vazio $S \in \boldsymbol{R}^n$ é chamado um subespaço de $\boldsymbol{R}^n$ se $\alpha x + \beta y \in S$ para todo $x, y \in S$ e $\alpha, \beta \in \boldsymbol{R}$.

Note que o vetor nulo sempre faz parte de um subespaço.

Além disso, se $S \neq \boldsymbol{R}^n$, dizemos que S é um subespaço próprio.

Subespaços e bases

O subespaço gerado por um número finito de vetores $x^1, ..., x^k \in \boldsymbol{R}^n$ é um subespaço formado por todos os vetores da forma $y = \sum_{i=1}^{k} \alpha_i x^i$, com α_i escalar.

Dado um subespaço S de $\boldsymbol{R}^n$, com $S \neq \{0\}$, uma base de S é uma coleção de vetores linearmente independente que gera S.

Toda base de um dado subespaço S de $\boldsymbol{R}^n$ tem o mesmo número de vetores.

Este número é chamado de dimensão de S.

Em particular, a dimensão de $\mathbb{R}^n$ é n e todo subespaço próprio de $\mathbb{R}^n$ tem dimensão menor do que n.

$\{0\}$ é um subespaço e sua dimensão é definida como 0.

Subespaços de dimensão 1 são retas que passam pela origem.

Subespaços de dimensão 2 são planos que passam pela origem.

Subespaços e bases

Se S é um subespaço próprio de $\boldsymbol{R}^n$, então existe um vetor não-nulo a que é ortogonal a S, ou seja, $a^T x = 0$ para todo $x \in S$.

De maneira mais geral, se S tem dimensão $m < n$, existem $n - m$ vetores linearmente independentes ortogonais a S.

Subespaços e bases

Teorema 2: *Suponha que o subespaço S gerado pelos vetores $x^1, ..., x^k$ tenha dimensão m. Então:*

(a) *Existe uma base de S composta com m dos vetores $x^1, ..., x^k$.*

(b) *Se $\ell \leq m$ e $x^1, ..., x^\ell$ são linearmente independentes, podemos formar uma base de S usando os vetores $x^1, ..., x^\ell$ e escolhendo $m - \ell$ vetores $x^{\ell+1}, ..., x^k$.*

Subespaços e bases

Seja A uma matriz de dimensão $m \times n$.

Chamamos de espaço coluna de A o subespaço de $\boldsymbol{R}^m$ gerado pelas n colunas de A.

Chamamos de espaço linha de A o subespaço de $\boldsymbol{R}^n$ gerado pelas m linhas de A.

Subespaços e bases

As dimensões dos espaços coluna e linha uma matriz A são sempre as mesmas.

Esta dimensão é chamada de posto de A. Claramente, $posto(A) \leq \min\{m, n\}$.

Dizemos que a matriz A tem posto completo se $posto(A) = \min\{m, n\}$.

O conjunto $\{x \in \boldsymbol{R}^n | Ax = 0\}$ é chamado de espaço nulo de A.

Ele é um subespaço de $\boldsymbol{R}^n$ e sua dimensão é $n - posto(A)$.

Subespaços afins

Sejam S_0 um subespaço de $\boldsymbol{R}^n$ e x^0 um vetor qualquer.

Se somarmos o vetor x^0 a cada vetor de S_0, estamos transladando S_0 de x^0.

O conjunto resultante pode ser definido como

$$S = S_0 + x^0 = \{x + x^0 | x \in S_0\}.$$

Este conjunto é chamado de subespaço afim.

Ele não necessariamente é um subespaço, já que pode não conter o vetor nulo.

A dimensão de S é definida como a dimensão de S_0.

Subespaços afins

Dados uma matriz $A \in \boldsymbol{R}^{m \times n}$ e um vetor $b \in \boldsymbol{R}^m$, considere o conjunto

$$S = \{x \in \boldsymbol{R}^n | Ax = b\},$$

que supomos ser não-vazio.

Vamos fixar um vetor x^0 tal que $Ax^0 = b$.

Um vetor x qualquer pertence a S se, e somente se,
$Ax = b = Ax^0 \Rightarrow A(x - x^0) = 0$.

Ou seja, x pertence a S se, e somente se, $x - x^0$ pertence ao subespaço $S_0 = \{y \in \boldsymbol{R}^n | Ay = 0\}$.

Portanto, $S = \{y + x^0 | y \in S_0\}$ e S é um subespaço afim de $\boldsymbol{R}^n$.

Subespaços afins

Se A tem m linhas linearmente independentes, seu espaço nulo S_0 tem dimensão $n - m$.

Portanto, o subespaço S também tem dimensão $n - m$.

Intuitivamente, se a_i são linhas de A, cada uma das restrições $a_i^T x = b_i$ remove um grau de liberdade de x, reduzindo a dimensão de n para $n - m$.