Autovalores e Autovetores

Considere uma matriz quadrada A_m, um escalar $r \in \mathbb{R}$ e um vetor $v \in \mathbb{R}^m$ com a seguinte propriedade:

$$A \cdot v = r \cdot v$$

Dizemos então que r é um *autovalor* e v é um *autovetor* da matriz A.

(Atenção: v é um vetor, e portanto não podemos 'cortar' v dos dois lados dessa equação: 'cortar' é uma operação de divisão, que na álgebra matricial corresponde a multiplicar por uma inversa, mas vetores não admitem inversa.)

Note que $v = (I \cdot v)$: a matriz identidade é o elemento neutro da multiplicação matricial. Logo, $A \cdot v = r(I \cdot v)$.

Portanto:

$$A \cdot v - r(I \cdot v) = 0$$

$$(A - r \cdot I)v = 0$$

Se $v \neq 0$, então a matriz $(A - r.I)$ não pode admitir inversa ($A - rI$ deve ser portanto uma matriz *singular*). Caso contrário, poderíamos multiplicar os dois lados da equação acima pela inversa $(A - r.I)^{-1}$, levando ao seguinte resultado absurdo:

$$0 \neq v = (A - r.I)^{-1} \cdot 0 = 0$$

(A última igualdade ocorre porque qualquer matriz multiplicada por um vetor de zeros gera um vetor de zeros).

Para que a matriz $A - rI$ não admita inversa, então necessariamente $det(A - r \cdot I) = 0$. Para calcular esse determinante, construa a matriz $A - rI$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix}$$

$$A - r.I = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} - r. \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$A - r.I = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{pmatrix} - \begin{pmatrix} r & 0 & \cdots & 0 \\ 0 & r & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & r \end{pmatrix}$$

$$A - r.I = \begin{pmatrix} a_{11} - r & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} - r & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} - r \end{pmatrix}$$

Ou seja: o autovalor $r \in \mathbb{R}$ é um número que quando subtraído da diagonal principal de uma matriz, a transforma em uma matriz singular.

Exemplo:

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

$r = 2 \to A - r.I = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow det = 0$

$r = 3 \to A - r.I = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow det = 0$

Logo, $r = 2$ e $r = 3$ são autovalores de A.

Os elementos da diagonal principal de uma matriz *diagonal* são os próprios autovalores: isso vale para qualquer matriz diagonal. Note ainda que $r = 0$ sempre é autovalor de uma matriz singular.

Exemplo:

$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ singular.

$r = 0 \to A - r.I = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ singular

Para encontrar os autovalores de uma matriz qualquer, usamos o polinômio característico dessa matriz. Considere uma matriz genérica de ordem 2:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$A - r.I = \begin{pmatrix} a_{11} - r & a_{12} \\ a_{21} & a_{22} - r \end{pmatrix}$$

$det(A - r.I) = 0$ é a condição para r ser autovalor:

$$(a_{11} - r)(a_{22} - r) - a_{12}a_{21} = 0$$

$$\underbrace{r^2 - (a_{11} + a_{22})r + (a_{11}a_{22} - a_{12}a_{21})}_{\text{polinomio caracteristico}} = 0$$

Uma equação de 2º grau tem até duas raízes distinas; logo, uma matriz 2x2 tem até dois autovalores distintos. Uma matriz $m\,x\,m$ tem até m autovalores distinos: as raízes do polinômio característico de ordem m.

Para encontrar o autovetor v^* associado ao autovalor r^*, basta resolver o sistema linear $(A - r^*I) \cdot v^* = 0$ para $v^* \neq 0$. Observe que esse é um sistema linear *homogêneo* com matriz de coeficientes singular (determinante é zero): logo, há infinitas soluções. Em outras palavras, há infinitos autovetores associados a um dado autovalor.

Exemplo:
(i) $A = \begin{pmatrix} -1 & 3 \\ 2 & 0 \end{pmatrix}$
$det(A - r.I) = det \begin{pmatrix} -1-r & 3 \\ 2 & -r \end{pmatrix} = r^2 + r - 6$: polinômio característico.
$r^2 + r - 6 = 0$: condição para r ser autovalor.
Raízes: $r_1 = -3$, $r_2 = 2$: autovalores (reais e distintos) de A.

Autovetor v_1 associado a r_1:
$(A - r_1 I) = \begin{pmatrix} -1-(-3) & 3 \\ 2 & -(-3) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
$\begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow 2x + 3y = 0 \Rightarrow y = -\frac{2}{3}x$: todos os pares (x, y) que respeitem essa equação são um autovetor da matriz A associado ao autovalor -3.
Em particular, $x = 1 \Rightarrow y = -\frac{2}{3}$: $v_1 = (1, -\frac{2}{3})$ é um autovetor associado a $r_1 = -3$. Qualquer múltiplo de v_1 é um autovetor associado a r_1 (há sempre infinitos autovetores associados a um dado autovalor).

Autovetor v_2 associado a r_2:
$(A - r_1 I)v = \begin{pmatrix} -1-2 & 3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
$\begin{pmatrix} -3 & 3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow -3x + 3y = 0 \Rightarrow y = x$
Em particular, $x = 1 \Rightarrow y = 1$: $v_2 = (1, 1)$ é um autovetor associado a $r_2 = 2$. Qualquer múltiplo de v_2 é um autovetor associado a r_2.

Conclusão: autevetor não é único.

Para concluir, defina o traço de uma matriz quadrada como a soma dos elementos da diagonal principal:
$Tr(A) = a_{11} + a_{22} + \cdots + a_{mm}$

Teorema: Considere a matriz quadrada A_m com autovalores $r_1, \ldots, r_m$.
(i) $r_1 + r_2 + \cdots + r_m = a_{11} + a_{22} + \cdots + a_{mm} = Tr(A)$
(ii) $r_1 \cdot r_2 \cdot \ldots \cdot r_m = det(A)$

Formas Quadráticas

Formas quadráticas são uma generalização das funções quadráticas univariadas $y = x^2$. Permitimos agora que o domínio seja $\mathbb{R}^n$.
Uma forma quadrática $Q : \mathbb{R}^n \longrightarrow \mathbb{R}$ admite representação matricial. Podemos defini-la da seguinte forma:

$$Q: \ \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$Q(x) = x^T A x$$

em que A é uma matriz quadrada de ordem n, e $x \in \mathbb{R}^n$ (ou seja, x é um vetor coluna com n coordenadas). Há infinitas matrizes que podem ser usadas para representar uma forma quadrática; escolheremos sempre uma matriz simétrica.

Exemplo: $Q : \mathbb{R} \longrightarrow \mathbb{R} : \ Q(x) = ax$. Como o domínio é unidimensional ($n = 1$), a matriz A tem ordem 1, ou seja, tem apenas uma linha e uma coluna: $A = [a]$. Podemos escrever $Q(x) = ax = Ax$. Observe que se $a > 0$, temos uma parábola 'virada para cima' que passa pelo ponto $(0,0)$; se $a < 0$, temos uma parábola 'virada para baixo'.

Exemplo: $Q : \ \mathbb{R}^2 \to \mathbb{R} : Q(x_1, x_2) = a_{11}x_1^2 + bx_1x_2 + a_{22}x_2^2$

$Q = \left[\begin{array}{cc} x_1 & x_2 \end{array} \right] \left[\begin{array}{cc} a_{11} & \frac{b}{2} \\ \frac{b}{2} & a_{22} \end{array} \right] \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right]$

Observe que podemos escolher outras matrizes para representar a forma quadrática acima. Considere uma matriz qualquer de ordem 2:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Obtemos então:

$$[x_1 \ x_2] \cdot A \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a_{11}x_1^2 + (a_{12} + a_{21})x_1 \cdot x_2 + a_{22}x_2^2$$

Logo, podemos escolher quaisquer elementos a_{12} e a_{21} tais que $a_{12} + a_{21} = b$. Como queremos usar uma matriz simétrica, fazemos $a_{12} = a_{21} = \frac{b}{2}$.

Definição de formas quadráticas

Note que uma forma quadrática não tem termo independente de x. Logo. $Q(0) = 0$ (observe que o 'zero' no domínio é, na verdade, a origem de $\mathbb{R}^n$: ou seja, é um vetor de zeros). Questão: $x = 0$ é um ponto de máximo, de mínimo, ou nenhum dos dois?

Em uma dimensão: $Q(x) = ax^2$. Se $a > 0$, $x = 0$ é um ponto de mínimo, e $Q > 0$ para todo $x \neq 0$. (represente graficamente essa função). Dizemos que a forma quadrática Q é positiva definida. Se $a < 0$, $x = 0$ é um ponto de máximo, e $Q < 0$ para todo $x \neq 0$. Dizemos que a forma quadrática Q é, nesse caso, negativa definida.

Podemos estender o raciocínio para duas ou mais dimensões. Se $(x_1, x_2, \ldots, x_n) \neq (0, \ldots, 0) \Rightarrow Q > 0$, então Q é positiva definida. Exemplo: $Q(x_1, x_2) = x_1^2 + x_2^2$

Se $(x_1, x_2, \ldots, x_n) \neq (0, \ldots, 0) \Rightarrow Q < 0$, então Q é negativa definida. Exemplo: $Q(x_1, x_2) = -x_1^2 - x_2^2$

Se uma função quadrática pode assumir valores positivos ou negativos, é dita indefinida. Exemplo: Exemplo: $Q(x_1, x_2) = x_1^2 - x_2^2$. Note que $Q(1, 0) > 0$ e $Q(0, 1) < 0$.

Se uma forma quadrática nunca assume valores negativos, mas pode ser igual a 0 para $x \neq 0$, é dita positiva semi-definida. Exemplo: $Q = x_1^2 + 2x_1x_2 + x_2^2 = (x_1 + x_2)^2$. Note que $Q(1, -1) = 0$.

Analogamente, a matriz negativa semi-definida nunca assume valores positivos, mas pode valer zero fora da origem. Exemplo: $-(x_1 + x_2)^2$

Podemos definir a matriz A a partir das propriedades da forma quadrática $Q(x) = x^T Ax, \ x \in \mathbb{R}^m$. Seja A_m simétrica. Então A é:

a) Positiva Definida se $x^T Ax > 0 \ \forall x \neq 0$.
b) Positiva Semi-Definida se $x^T Ax \geq 0 \ \forall x \neq 0$.
c) Negativa Definida se $x^T Ax < 0 \ \forall x \neq 0$.
d) Negativa Semi-Definida se $x^T Ax \leq 0 \ \forall x \neq 0$.
e) Indefinida caso não se encaixe em nenhum dos casos anteriores.

Teste para Definição de Matriz

Relembrando: *menores* são os determinantes de *submatrizes* formadas a partir da eliminação de determinadas linhas e colunas de uma matriz A. Se forem retiradas linhas e colunas de mesmo índice, temos um *menor principal* (por exemplo, se excluímos a linha 1 e a coluna 1).

Exemplo: A = $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

Menor principal de ordem 2: Det(A)

Menores principais de ordem 1: $Det(a_{11}) = a_{11}$ (determinante da submatriz formada pela eliminação da linha 2 e da coluna 2) e $Det(a_{22}) = a_{22}$ determinante da submatriz formada pela eliminação da linha 1 e da coluna 1

Exemplo: $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Menor principal de ordem 3: $Det(A)$.

Menores principais de ordem 2:

$det\left(\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}\right)$, obtido pela eliminação da linha 1 e da coluna 1;

$det\left(\begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}\right)$, obtido pela eliminação da linha 2 e da coluna 2;

$det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right)$, obtido pela eliminação da linha 3 e da coluna 3.

Menores principais de ordem 1:

$det(a_{33})$, obtido pela eliminação das linhas/colunas 1 e 2;

$det(a_{22})$, obtido pela eliminação das linhas/colunas 1 e 3;

$det(a_{11})$, obtido pela eliminação das linhas/colunas 2 e 3.

Os menores principais mais usados , para cada ordem, são formados por eliminação das *últimas* linhas e colunas. São conhecidos como Menores Principais Líderes (MPL's):

a_{11}, $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Teorema: Seja A uma matriz quadrada simétrica de ordem n. Então:

1 - A é positiva definida se e só se todos os MPL's são estritamente positivos.

2 - A é negativa definida se e só se os MPL's altenarem de sinal, começando com o sinal negativo: $A_1 < 0, A_2 > 0, ...$

Exemplos:

1 - $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ $\quad A_1 = 2 > 0,\ A_2 = 2 \times 2 - 1 \times 1 = 3 > 0:$ P.D

2 - $\begin{bmatrix} -1 & -2 \\ 2 & 1 \end{bmatrix}$ $\quad A_1 = -1 < 0, A_2 = (-1) \times 1 - 2 \times (-2) = 3 > 0:$ N.D

3 - $\begin{bmatrix} 2 & 3 \\ 3 & 7 \end{bmatrix}$ $\quad A_1 = 2 > 0, A_2 = 2 \times 7 - 3 \times 3 = 5 > 0:$ P.D

4 - $Q = x_1^2 + x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$A_1 = 1 > 0,\ \ A_2 = 1 \times 1 - 0 \times 0 = 1 > 0:\quad$ P.D.

5 - $Q = -x_1^2 - x_2^2 = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \underbrace{\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 & -x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$A_1 = -1 < 0, A_2 = (-1) \times (-1) - 0 \times 0 = 1 > 0$: N.D

Para checar se a matriz é semi-definida , são necessários todos os menores principais, não apenas os líderes.

Matrizes Diagonais

Considere uma matriz diagonal qualquer:

$$A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_n \end{bmatrix}$$

Essa matriz representa a forma quadrática $Q = a_1 x_1^2 + \cdots + a_n x_n^2$, em que não há produtos cruzados. Observe então que A é:

i- positiva definida se $a_i > 0\ \forall i$.

ii- positiva semi-definida se $a_i \geq 0\ \forall i$.

iii- negativa definida se $a_i < 0\ \forall i$.

iv- negativa semi-definida se $a_i \leq 0\ \forall i$.

v- indefinida se existe $a_i > 0$ e $a_j < 0$.

Note que, para uma matriz positiva definida, temos: $a_1 > 0$; $a_1 \cdot a_2 > 0$; $a_1 \cdot a_2 \cdot a_3 > 0, ...$: o produto de números positivos é sempre positivo. Já para a matriz negativa definida, temos: $a_1 < 0$; $a_1 \cdot a_2 > 0$; $a_1 \cdot a_2 \cdot a_3 < 0, ...$: o produto é negativo quando multiplicamos um número ímpar de números negativos, mas é positivo quando multiplicamos um número par de termos (o produto de dois números negativos é positivo). O teorema anterior generaliza essa relação para matrizes não-diagonais.

Para uma matriz diagonal, os elementos da diagonal principal são exatamente seus autovalores. Logo, uma matriz diagonal é positiva definida se e só se todos os autovalores são estritamente positivos, e positiva semi-definida caso sejam maiores ou iguais a zero. Analogamente, será negativa definida se e só se todos os autovalores forem estritamente negativos.

É possível provar que esse resultado vale para quaisquer matrizes, e não apenas para as diagonais: podemos avaliar a definição de uma forma quadrática a partir dos autovalores da matriz que a representa.

Equações e Sistemas de Diferença

Equações de Diferença

Um exemplo de equação de diferença é:

$y_{n+1} = a \cdot y_n$

Em que a é uma constante qualquer. O valor de y em um período é igual ao seu valor no período anterior multiplicado por essa constante.

Essa equação descreve, por exemplo, o retorno de um investimento sem risco. Ao investir uma quantia y_n no período n, obtemos um valor de juros de $\rho \cdot y_n$. Temos então no próximo período:

$$y_{n+1} = y_n + \rho \cdot y_n = (1 + \rho) y_n \tag{1}$$

Fazendo $1 + \rho = a$, obtemos a equação anterior.

Para resolver essa equação, precisamos conhecer o valor inicial y_0. Então:

$$\begin{aligned} y_1 &= a y_0 \\ y_2 &= a y_1 = a(a y_0) = a^2 y_0 \\ y_3 &= a y_2 = a(a^2 y_0) = a^3 y_0 \end{aligned} \tag{2}$$

E assim em diante. A solução geral é $y_n = a^n y_0$. Se $|a| < 1$, então o sistema é *estável*: converge para zero quando n aumenta.

Sistema de Diferenças: Duas Equações

Considere agora o sistema formado pelas duas equações abaixo:

$$\begin{aligned} x_{n+1} &= a \cdot x_n + b \cdot y_n \\ y_{n+1} &= c \cdot x_n + d \cdot y_n \end{aligned} \tag{3}$$

Esse sistema pode ser representado em forma matricial. Defina:

$$z_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix} \tag{4}$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{5}$$

Podemos então escrever:

$$z_{n+1} = A \cdot z_n \tag{6}$$

Para verificar que essa representação está de acordo com o sistema original, basta fazer a multiplicação:

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} ax_n + by_n \\ cx_n + dy_n \end{bmatrix} \tag{7}$$

Se $b = c = 0$, as equações não são relacionadas,e temos apenas:

$$\begin{aligned} x_{n+1} &= ax_n \\ y_{n+1} &= dy_n \end{aligned} \tag{8}$$

As soluções já foram estabelecidas nesse caso: $x_n = a^n x_0$ e $y_n = a^n y_0$. *Esse é o caso geral para matrizes diagonais.*

Para resolver o caso geral, em que as equações podem ser relacionadas (ou seja, $b \neq 0$ ou $c \neq 0$), precisamos considerar um vetor inicial $z_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$.

Temos então:

$$\begin{aligned} z_1 &= Az_0 \\ z_2 = Az_1 &= A^2 z_0 \\ &\vdots \\ z_n &= A^n z_0 \end{aligned} \tag{9}$$

Precisamos então computar A^n. Se $A = D$ é uma matriz diagonal, essa conta é fácil:

$$D \cdot D = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \cdot \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} = \begin{bmatrix} {r_1}^2 & 0 \\ 0 & {r_2}^2 \end{bmatrix} \tag{10}$$

Podemos estender esse raciocínio:

$$D^n = \begin{bmatrix} {r_1}^n & 0 \\ 0 & {r_2}^n \end{bmatrix} \tag{11}$$

Precisamos então de uma matriz não-singular P tal que $P^{-1}AP$ seja uma matriz diagonal, que denotaremos por D. Nesse caso, $A = PDP^{-1}$, e portanto:

$$\begin{aligned} A^2 &= (PDP^{-1})(PDP^{-1}) = (PD)(P^{-1}P)(DP^{-1}) \\ &= (PD)I(DP^{-1}) = (PD)(DP^{-1}) = P(D^2)P^{-1} \end{aligned} \tag{12}$$

De forma geral:

$$A^n = P(D^n)P^{-1} \tag{13}$$

Ou seja:

$$A^n = P \begin{bmatrix} {r_1}^n & 0 \\ 0 & {r_2}^n \end{bmatrix} P^{-1} \tag{14}$$

A solução do sistema é portanto:

$$z_n = \underbrace{[P \begin{bmatrix} {r_1}^n & 0 \\ 0 & {r_2}^n \end{bmatrix} P^{-1}]}_{A^n} z_0 \tag{15}$$

Para obter P e D, escreva $P = [v_1 \ v_2]$, em que v_1 e v_2 representam as colunas de P. $P^{-1}AP = D$ pode ser reescrito como $AP = PD$. Ou seja:

$$A[v_1 \ v_2] = [v_1 \ v_2] \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \tag{16}$$

Ou ainda:

$$\begin{bmatrix} Av_1 & Av_2 \end{bmatrix} = \begin{bmatrix} r_1 v_1 & r_2 v_2 \end{bmatrix} \tag{17}$$

Logo, $Av_1 = r_1 v_1$ e $Av_2 = r_2 v_2$. Mas isso significa que r_1 e r_2 são autovalores de A, e v_1 e v_2 são os autovetores associados.

Para construir a solução explicitamente:

$$\begin{aligned} z_n &= A^n z_0 \\ z_n &= PD^n P^{-1} z_0 \\ P^{-1} z_n &= P^{-1} PD^n P^{-1} z_0 \\ P^{-1} z_n &= ID^n P^{-1} z_0 \\ P^{-1} z_n &= D^n P^{-1} z_0 \end{aligned} \tag{18}$$

Defina agora $Z_n = P^{-1} z_n = P^{-1} \begin{bmatrix} x_n \\ y_n \end{bmatrix}$. Logo, $Z_n = D_n Z_0$. Como D^n é diagonal, conseguimos encontrar a solução. Para tanto, defina X_n como a primeira linha de Z_n, e Y_n como a segunda linha:

$$\begin{aligned} X_n &= c_1 r_1^n \\ Y_n &= c_2 r_2^n \end{aligned} \tag{19}$$

Em que c_1 e c_2 são as condições iniciais das variáveis artificiais X_n e Y_n, respectivamente (se $n = 0$, $X_n = c_1$, pois $r_1^0 = 1$).

Para recuperar z_n, fazemos:

$$z_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix} = P \begin{bmatrix} X_n \\ Y_n \end{bmatrix} = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \begin{bmatrix} c_1 r_1^n \\ c_2 r_2^n \end{bmatrix} = c_1 r_1^n v_1 + c_2 r_2^n v_2 \tag{20}$$

Ou seja, temos uma combinação linear dos autovetores de A em que os coeficientes dependem das condições iniciais e dos autovalores correspondentes.

Observe então que o sistema é estável se ambos os autovalores são menores do que um em módulo: nesse caso, r_i^n se aproxima de zero quando n aumenta.

Considere o seguinte exemplo:

$$\begin{aligned} x_{n+1} &= x_n + 4y_n \\ y_{n+1} &= 0.5x_n \end{aligned} \tag{21}$$

Temos então a seguinte matriz de coeficientes:

$$A = \begin{bmatrix} 1 & 4 \\ 0.5 & 0 \end{bmatrix}$$

Os autovalores dessa matriz são 2 e -1. Os autovetores correspondentes são $(1, \frac{1}{4})$ e $(1, \frac{-1}{2})$.

A solução é portanto:

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = c_1 \cdot 2^n \cdot \begin{bmatrix} 1 \\ \frac{1}{4} \end{bmatrix} + c_2 \times (-1)^n \cdot \begin{bmatrix} 1 \\ \frac{-1}{2} \end{bmatrix} \tag{22}$$

Em que c_1 e c_2 dependem das condições iniciais, como veremos abaixo.

Ou seja:

$$\begin{aligned} x_n &= c_1 \cdot 2^n + c_2 \cdot (-1)^n \\ y_n &= \frac{c_1 \cdot 2^n}{4} - \frac{c_2 \cdot (-1)^n}{2} \end{aligned} \tag{23}$$

Para determinar as contantes c_1 e c_2, precisamos do vetor inicial (x_0, y_0). Fazendo $n = 0$ nas equações acima, obtemos:

$$\begin{aligned} x_0 &= c_1 + c_2 \\ y_0 &= \frac{c_1}{4} - \frac{c_2}{2} \end{aligned} \tag{24}$$

Resolvendo esse sistema, obtemos:

$$\begin{aligned} c_1 &= \frac{x_0}{6} + \frac{y_0}{3} \\ c_2 &= \frac{-x_0}{6} + \frac{2y_0}{3} \end{aligned} \tag{25}$$

A solução fica então inteiramente caracterizada. Se, por exemplo, $x_0 = 2$ e $y_0 = -\frac{1}{4}$, encontramos $c_1 = c_2 = 1$, e portanto a solução do sistema é:

$$\begin{aligned} x_n &= 2^n + (-1)^n \\ y_n &= \frac{2^n}{4} - \frac{(-1)^n}{2} \end{aligned} \tag{26}$$

Observe que o sistema não é estável porque os autovalores não são inferiores a 1 em módulo.

Processos de Markov

Considere uma variável S que pode assumir k valores: $S_1, S_2, ..., S_k$. Cada um desses valores é chamado *estado*. Por exemplo, um indivíduo no mercado de trabalho pode ser classificado como empregado ou desempregado.

Definição: um *Processo Estocástico* é uma regra que determina a probabilidade de que uma variável esteja no estado i no período $n+1$, dadas as probabilidades de que estivesse em cada um dos k estados possíveis nos períodos anteriores $(1, 2, ..., n)$.

Definição: um *Processo de Markov* é um processo estocástico em que a probabilidade de cada estado no período $n+1$ depende apenas do estado em n: só o passado imediato importa.

Os objetos centrais para descrever um processo de Markov são:

1- As probabilidades $x_i(n)$ de que o estado i ocorra no período n (interpretado como a fração da população total nesse estado, nesse período);

2- Probabilidades de transição: $m_{ij} = Prob[S_i(n+1) \text{ dado } S_j(n)]$.

Podemos fazer uma representação matricial das probabilidades de transição:

$$M = \begin{bmatrix} m_{11} & \dots & m_{1k} \\ \vdots & & \vdots \\ m_{k1} & \dots & m_{kk} \end{bmatrix}$$

Essa á chamada matriz de transição, ou estocástica, ou ainda matriz de Markov. Observe que pela definição de m_{ij}, o primeiro subscrito indica o estado no *próximo período*, e o segundo subscrito indica o estado *no período atual*. Logo, a soma de cada coluna deve ser igual a um: há 100% de chance de estar em *algum* estado, em cada período.

As probabilidades m_{ij} são fixas, independentes do período n. São ditas estacionárias.

Se N é a população total e $x^j(n)$ é a fração dessa população no estado j no período n (por exemplo, taxa de desemprego), o número de pessoas nesse estado, nesse período, é $x^j(n) \cdot N$ (número de desempregados).

Por construção, $m_{ij} \cdot x^j(n) \cdot N$ indivíduos estarão no estado i, no período $n+1$. Logo, o número total de indivíduos no estado i, em $n+1$, será:

$$\sum_{j=1}^{k} m_{ij} \cdot x^j(n) \cdot N$$

Por exemplo: o número total de desempregados no período $n+1$ é a soma de (i) desempregados em n que continuaram desempregados; (ii) empregados em n que perderam o emprego.

Novamente, podemos usar uma notação matricial.

$$\begin{bmatrix} x^1(n+1) \\ \vdots \\ x^k(n+1) \end{bmatrix} = \begin{bmatrix} m_{11} & \dots & m_{1k} \\ \vdots & & \vdots \\ m_{k1} & \dots & m_{kk} \end{bmatrix} \begin{bmatrix} x^1(n) \\ \vdots \\ x^k(n) \end{bmatrix}$$

Ou seja: $x(n+1) = M \cdot x(n)$. Trata-se de um sistema de diferenças.

Exemplos: economia do trabalho: há apenas dois estados, empregado (com fração $x^1(n)$) ou desempregado (com fração $x^2(n)$).

$$\begin{bmatrix} x^1(n+1) \\ x^2(n+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 0.9 & 0.4 \\ 0.1 & 0.6 \end{bmatrix}}_{\text{M}} \begin{bmatrix} x^1(n) \\ x^2(n) \end{bmatrix}$$

Interpretação: há 90% de chance de um empregado continuar empregado, e 10% de chance de perder o emprego; há 60% de chance de um desempregado continuar desempregado, e 40% de chance de conseguir emprego.

Observe que $r_1 = 1$ é autovalor de M (o que é sempre verdade para matrizes de Markov). Subtraindo de cada termo da diagonal principal de M, obtemos:

$$M - 1 \cdot I = \begin{bmatrix} -0.1 & 0.4 \\ 0.1 & -0.4 \end{bmatrix}$$

Essa matriz é singular (o determinante é igual a zero). Use agora a propriedade de que o traço é igual à soma dos termos da diagonal principal de M:

$$a_{11} + a_{22} = r_1 + r_2$$

$$0.9 + 0.6 = 1 + r_2$$

Logo, $r_2 = 0.5$.

Podemos agora obter os autovalores. Para $r_1 = 1$:

$$\begin{bmatrix} -0.1 & 0.4 \\ 0.1 & -0.4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longrightarrow 0.1 \cdot x = 0.4 \cdot y \Longrightarrow x = 4y$$

Logo, o autovalor é $(4, 1)$. Analogamente, obtemos para $r_2 = 0.5$:

$$\begin{bmatrix} 0.4 & 0.4 \\ 0.1 & 0.1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Longrightarrow x = y$$

Logo, o autovetor associado a $r_2 = 0.5$ é $(1, -1)$.

Podemos então escrever a solução geral do sistema de diferenças:

$$\begin{bmatrix} x^1(n) \\ x^2(n) \end{bmatrix} = c_1 \begin{bmatrix} 4 \\ 1 \end{bmatrix} \cdot 1^n + c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot (0.5)^n$$

Como $|0.5| < 1$, concluímos que quando $n \longrightarrow \infty$, o sistema converge para:

$$\begin{bmatrix} x^1(n) \\ x^2(n) \end{bmatrix} \longrightarrow c_1 \cdot \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 4c_1 \\ c_1 \end{bmatrix}$$

Como estamos trabalhando com um vetor de probabilidades, que devem somar um, podemos fazer $4c_1 + c_1 = 1 \Longrightarrow c_1 = 0.2$. Concluímos então que o sistema converge para:

$$(0.8\ 0.2)$$

Ou seja, o desemprego de longo prazo é 20%.

O exemplo anterior usa um matriz de transição *positiva*: todos os elementos são estritamente positivos. De forma mais geral, dizemos que uma matriz M de Markov é *regular* se existe um número inteiro r tal que M^r é positiva. Podemos provar para matrizes regulares:

a- $r_1 = 1$ é autovalor de M;

b- os demais autovalores são menores que um em módeulo;

c- todas as entradas do autovetor w_1 associado ao autovalor $r_1 = 1$ são estritamente positivas;

d- se v_1 é o vetor w_1 dividido pela soma das suas coordenadas, concluímos que $x(n) \longrightarrow v_1$ quando $n \longrightarrow \infty$.

Se existe $m_{ii} = 1$, então S_i é dito estado absorvente: uma vez atingido, o sistema não muda mais.

Matrizes Simétricas

Grande parte das matrizes em otimização e econometria (principais aplicações de Álgebra Linear em economia) são simétricas: Hessiana, matriz de variância-covariância, matriz de informação, entre outras. Temos o seguinte resultado:

Teorema: Se A é simétrica de ordem k, então:

a- Todos os autovalores são reais;

b- Autovetores associados a diferentes autovalores são ortogonais;

c- Existe uma matriz P não-singular, cujas colunas $w_1, ..., w_k$ são os autovetores de A, tal que: (i) $w_1, ..., s_k$ são ortogonais; (ii) $P{-}1 = P^T$; (iii) $P^{-1}AP = P^TAP$ é uma matriz diagonal cujos elementos (na diagonal principal) são os autovalores de A. (Esse resultado vale mesmo que A tenha autovalores repetidos.)

Dizemos que P é uma matriz ortonormal: $P^{-1} = P^T$.

Espaço Vetorial (SB 27)

O espaço euclidiano $\mathbb{R}^n$ é um espaço vetorial porque admite duas operações naturais: soma e multiplicação por escalar, que respeitam uma série de propriedades, descritas abaixo.

Seja V um espaço vetorial qualquer. Sejam u, v, w elementos desse espaço: $u, v, w \in V$. Sejam r e s escalares: $r, s, \in \mathbb{R}$. As seguintes propriedades devem valer para que V seja um E.V:

$(i)\, u + v \in V$: V é fechado para adição vetorial.
$(ii)\, u + v = v + u$
$(iii)\, u + (v + w) = (u + v) + w$
(iv) Existe um elemento nulo 0 / $v + 0 = v$
(v) Para todo v, existe um elemento $-v \in V/v + (-v) = 0$
(vi) Para todo escalar $r, r \cdot v \in V$: V é fechado para multiplicação por escalar.
$(vii)\, r \cdot (u + v) = r.u + r.v$
$(viii)\, (r + s) \cdot u = r \cdot u + s \cdot u$
$(ix)\, r \cdot (s \cdot u) = (r \cdot s) \cdot u$
$(x)\, 1 \cdot u = u$

Exemplos de espaços vetoriais (além de $\mathbb{R}^n$):

- Conjunto das sequências infinitas de números reais, com as operações de soma e multiplicação por escalar iguais às definidas para vetores com um número qualquer (finito) de elementos;

- Conjunto $M(m,n)$ de todas as matrizes com m linhas e n colunas, com as operações já definidas de soma e de multiplicação por escalar;

Um sub-espaço vetorial é um subconjunto $V_0 \subset V$ que também possui as mesmas propriedades acima.

Na prática: considere um conjunto $V_0 \subset V$, em que V é espaço vetorial. Se V_0 for fechado para adição e para multiplicação por escalar, e contiver a origem, então é um subespaço vetorial de V.

Ou seja: se $V_0 \subset V$ for tal que todas as combinações lineares $r \cdot v + s \cdot u \in V_0$ para quaisquer $v, u \in V_0$, então V_0 é subespaço vetorial de V.

Exemplos:

i- Seja $w \in V$. O conjunto de todos os múltiplos de w - ou seja, a reta que passa pela origem e contém w - é subespaço vetorial: $\mathcal{L}(w) = \{\alpha \cdot w, \alpha \in \mathbb{R}\}$.

ii- (Caso particular do item i.) Considere o vetor $w = (1,1,1) \in \mathbb{R}^3$. O conjunto $V_0 = \mathcal{L}(w) = \{r \cdot (1,1,1),\ r \in \mathbb{R}\} = \{(r,r,r),\ r \in \mathbb{R}\}$ é subespaço vetorial de $\mathbb{R}^3$.

iii- (Generalização do item i.) Sejam $w_1, \ldots, w_k$ vetores em $\mathbb{R}^n$. $V_0 = \mathcal{L}(w_1, \ldots, w_k) = \{r_1 w_1, \ldots, r_k w_k\ :\ r_1, \ldots, r_k \in \mathbb{R}\}$: o conjunto de combinações lineares de $w_1, \ldots, w_k$ é subespaço vetorial.

iv- $V_0 = \{(x_1, x_2, 0),\ x_1, x_2 \in \mathbb{R}\}$, $V = \mathbb{R}^3$

v- $V_0 = \{(x_1, x_2, 1) : x_1, x_2 \in \mathbb{R}\}$ não é subespaço, pois não passa pela origem: não inclui o ponto $(0,0,0)$, já que a terceira coordenada é sempre igual a 1.

vi-$V_0 = \{0\}$ é subespaço vetorial de V.

vii- $V_0 = V$ é subespaço vetorial de V.

Os dois últimos exemplos são os chamados subespaços triviais de V: o conjunto formado apenas pela origem, e o espaço inteiro.

Se m vetores $u_1, \ldots, u_m$ formam base do subespaço vetorial V_0, então qualquer subconjunto com mais de m vetores é linearmente dependente (LD). Qualquer base de V_0 tem o mesmo número m de vetores, e esse número é a dimensão do sub-espaço V.

Espaço-linha de uma matriz

Considere uma matriz $A_{(n\,x\,m)}$. Cada uma das n linhas de A pode ser vista como um vetor em $\mathbb{R}^m$.

Considere o espaço gerado por esses vetores (ou seja, pelas n as linhas de A): $Row(A) = \mathcal{L}(a_1, \ldots, a_n)$

$a_1, \ldots, a_n$ são as linhas de A. $\mathcal{L}(a_1, \ldots, a_n)$ é o conjunto de C.L's dessas linhas. Operações elementares com linhas mudam as linhas $a_1, \ldots, a_n$, mas não mudam o espaço gerado por elas!

Se $\widetilde{a}_1, \ldots, \tilde{a}_n$ são linhas da matriz $\widetilde{A}$ gerada a partir de operações elementares sobre A, então: $\mathcal{L}(a_1, \ldots, a_n) = \mathcal{L}(\widetilde{a}_1, \ldots, \tilde{a}_n)$. Em particular, isso vale para a forma escalonada de A.

Se os espaços gerados são os mesmos, então a dimensão deles é sempre a mesma. A dimensão não depende das operações elementares efetivadas sobre as linhas de A.

O posto de A é a dimensão do espaço gerado pelas linhas de A, e portanto é invariante a operações elementares com linhas.

Espaço Coluna

$A_{(n\,x\,m)}$: m vetores em $\mathbb{R}^n$: $a_1, \ldots, a_m$.
Espaço coluna de A: $\mathcal{L}(a_1, \ldots, a_m)$

Definição: uma coluna é dita *básica* se a coluna correspondente na matriz escalonada tiver um pivô.

Resultado: as colunas básicas de A forma um subespaço para o espaço coluna de A.

Então:

$$\begin{aligned} Dim(E.L) &= \#linhas \neq 0 \text{ na forma escalonada} \\ &= \#\text{pivôs na forma escalonada} \\ &= Dim(E.C) \end{aligned}$$

Então: $Dim(E.C) = Dim(E.L) = posto$

Sabemos que $Ax = b$ possui solução se e somente se $b \in E.C(A)$:

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_m \begin{pmatrix} a_{1m} \\ \vdots \\ a_{mm} \end{pmatrix}$$

De forma equivalente:
$x_1 a_1 + \cdots + x_m a_m = b$
O lado esquerdo dessa equação é uma combinação linear das colunas da matriz A.

Logo, $Ax = b$ possui solução $\forall b \in \mathbb{R}^m$ se e somente se $E.C(A) = \mathbb{R}^n$; i.e, a dimensão de $E.C(A)$ é m.

Espaço Nulo

O conjunto V de todas as soluções do S.L $Ax = b$ é um espaço vetorial. É chamado espaço nulo de A: $N(A)$.
Um *espaço afim* é um deslocamento paralelo de um E.V: ou seja, é o conjunto $\{x \in \mathbb{R}^m : x = c + v \ , \, v \epsilon \ V\}$.
Espaços-afins são soluções para sistemas não-homogêneos $Ax = b \neq 0$.

Teorema Fundamental da Álgebra Linear

Qual a relação entre $Dim(N(A))$ e $Posto(A)$?

$Posto(A) + Dim(N(A)) = \#colunas$ de A

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} : \ Posto(A) = 2 = \#linhas \to Dim(N(A)) = 0 : \ \exists! \text{ solução.}$$

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} : \ Posto(A) = 1, \ \#linhas = 2 \to Dim(N(A)) = 1 : \text{ reta.}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$: $Posto(A) = 1, \#linhas = 3 \rightarrow Dim(N(A)) = 2$: plano.

Transformações Lineares

Sejam V e W espaços vetoriais. A função $T : V \longrightarrow W$ é uma transformação linear se, para quaisquer $v_1, v_2 \in V$, e para qualquer escalar $c \in \mathbb{R}$, temos:

1- $T(v_1 + v_2) = T(v_1) + T(v_2)$

2- $T(c \cdot v_1) = c \cdot T(v_1)$

De forma equivalente, podemos dizer que para quaisquer $v_1, v_2 \in V$ e $c_1, c_2 \in \mathbb{R}$,

$$T(c_1v_1 + c_2v_2) = c_1T(v_1) + c_2T(v_2)$$

Ou seja, a transformação linear de uma combinação linear é a combinação linear das transformações dos vetores individuais.

Observe que essa definição implica que uma transformação linear sempre passa pela origem:

$$T(0) = 0$$

A demonstração é simples:

$T(0) = T(0 + 0) = T(0) + T(0)$. Subtraindo $T(0)$ de ambos os lados, obtemos $T(0) - T(0) = T(0) + T(0) - T(0)$, ou $0 = T(0)$, o que prova o resultado.

Portanto, a função $y = ax + b$ não é uma transformação linear se $b \neq 0$, pois não passa no ponto $(0, 0)$.

Exemplos

1- Seja $T : \mathbb{R} \longrightarrow \mathbb{R}$ definida por $T(x) = a \cdot x$ para algum número real a. Podemos verificar que essa função define uma transformação linear:

$$T(x_1 + x + 2) = a(x_1 + x_2) = ax_1 + ax_2 = T(x_1) + t(x_2)$$

$$T(cx) = a \cdot (cx) = c \cdot (ax) = c \cdot T(x)$$

Logo, as duas condições são atendidas.

2- $T : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = a_1x_1 + a_2x_2 = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

3- $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida por:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right) = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

4- $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \\ a_{31}x_1 + a_{32}x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

5- $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ definida por:

$$T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} a_{11}x_1 + ... + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + ... + a_{mn}x_m \end{bmatrix} = \underbrace{\begin{bmatrix} a_{11} & ... & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & ... & a_{mn} \end{bmatrix}}_{A_{m\times n}} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Ou seja:

$$T\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = A \cdot \left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)$$

Toda transformação linear (finita) pode ser representada dessa forma.

Complemento e Projeção Ortogonal

Complemento Ortogonal

Considere um subespaço $V \subseteq E$. Considere $E = \mathbb{R}^n$. Defina o complemento ortogonal de V:

$$V^{\perp} = \{x \in \mathbb{R}^n :< x, v >= 0 \; \forall \, v \in V\}$$

Qualquer elemento de $V^{\perp}$ é perpedicular a qualquer elemento de V. Como exemplo, podemos tomar $E = \mathbb{R}^2$, V como o conjunto de múltiplos de $(1, 0)$ (eixo horizontal), e portanto $V^{\perp}$ como o conjunto de múltiplos de $(0, 1)$ (eixo vertical).

É possível provar os seguintes resultados.

Resultado 1. *$V^{\perp}$ também é subespaço vetorial de E.*

Prova. *Vamos provar inicialmente que $V^{\perp}$ é fechado para soma. Sejam $a, b \in V^{\perp}$. Precisamos mostrar que $a + b \in V^{\perp}$, ou seja, precisamos provar que $a + b$ é perpendicular a qualquer elemento $v \in V^{\perp}$. Mas isso equivale a provar que o produto interno $< a+b, v >$ é igual a zero para todo $v \in V^{\perp}$. Pelas propriedades do produto interno, $< a+b, v >=< a, v > + < b, v >$. Como $a \in V^{\perp}$, a é perpendicular a qualquer elemento $v \in V^{\perp}$, e portanto $< a, v >= 0$. Analogamente, $< b, c >= 0$, e portanto $< a + b, v >= 0$, e que conclui a prova de que $V^{\perp}$ é fechado para soma.*

Para provar que $V^{\perp}$ é fechado para multiplicação por escalar, precisamos mostrar que para todo $a \in V^{\perp}$ e para todo $\alpha in \mathbb{R}$, $\alpha \cdot a \in V^{\perp}$. Ou seja, temos de provar que $< \alpha a, v >= 0$ para todo $v \in V$. Usamos então novamente as propreidades do produto interno: $< \alpha a, v >= \alpha < a, v >= 0$, em que a última igualdade decorre do fato de que $a \in V^{\perp}$, e portanto

$\alpha < a, v >= 0$. Por último, podemos fazer $\alpha = 0$, e portanto $\alpha a = 0$: a origem também pertece ao conjunto $V^\perp$. Isso conclui a demonstração: $V^\perp$ é fechado para soma de vetores e para multiplicação por escalar, e contém a origem.

Resultado 2. *O núcleo de uma matriz A (ou seja, seu espaço-nulo) é igual ao conjunto ortogonal ao espaço-linha dessa matriz: $N(A) = (E.L.(A))^\perp$.*

Analogamente, podemos escrever o segundo resultado como $N(A) = (E.C.(A^T))^\perp$, pois espaço-linha de A é por definição o espaço-coluna de A^T, a transposta de A. Podemos ainda escrever:

$$N(A^T) = (E.L.(A^T))^\perp$$

Dimensão do Complemento Ortogonal

Podemos então determinar a dimensão de $V^\perp$, o complemento ortogonal de V.

Resultado 3. *Seja V subespaço vetorial de E, com $dim(V) = k$ e $dim(E) = n \geq k$. Então $dim^\perp = n - k$.*

Prova. *Seja $\{v_1, ..., v_k\}$ uma base para V, que tem portanto dimensão $dim(V) = k$ (lembrando: dimensão é o número de vetores em uma base).*

Construa uma matriz em que cada coluna é um vetor dessa base de V:

$$A_{n,k} = \begin{bmatrix} | & ... & | \\ v_1 & ... & v_k \\ | & ... & | \end{bmatrix}$$

Logo, $V = Span(v_1, ..., v_k) = E.C.(A)$. Lembrando: V é por definição o espaço gerado (span) dos vetores de uma base de V. Pela disposição desses vetores na matriz acima, isso é por definição o espaço-coluna de A.

Podemos então concluir que $N(A^\perp) = (E.C.(A))^\perp = V^\perp$. Logo, a dimensão de $V^\perp$ é igual à de $N(A^T)$. Escreva agora a transposta de A:

$$A^T_{k,n} = \begin{bmatrix} \text{——}v_1\text{——} \\ \vdots \\ \text{——}v_k\text{——} \end{bmatrix}$$

O teorema fundamentl da álgebra linear diz que:

$$Posto(A^T) + dim(Nucleo(A^T)) = \#\ colunas\ (A)$$

Mas o número de colunas de A^T é n. Além disso, o posto de A^T é igual à dimensão do espaço-linha de A^T, e o espaço-linha de A^T é o espaço-coluna de A. Logo, $posto(A^T) = dim(E.C.(A))$.

$$dim(E.C.(A)) + dim(Nucleo(A^T)) = n$$

Mas o espaço-coluna de A é por definição o subespaço V (os vetores da base de V foram dispostos como colunas da matriz A). E, como visto acima, $dim(V^{\perp}) = dim(N(A^T))$. Reescrevemos então a equação acima como:

$$dim(V) + dim(V^{\perp}) = n$$

Ou ainda:

$$dim(V^{\perp}) = n - dim(V)$$

É possível provar que qualquer vetor em $\mathbb{R}^n$ pode ser representado, *de forma única*, como a soma de um vetor $v \in V$ com um vetor $w \in V^{\perp}$. Por exemplo, qualquer vetor (x, y) em $\mathbb{R}^2$ pode ser escrito como a soma do vetor $(x, 0)$ no eixo horizontal com o vetor $(0, y)$ no eixo vertical (ou seja, o complemento ortogonal do eixo horizontal).

Decomposição de um vetor em $\mathbb{R}^n$ como uma soma de elementos em V e em $V^{\perp}$

Seja $\tilde{v} \in \mathbb{R}^n$. Queremos escrever $\tilde{v} = v + w$, com $v \in V$ e $w \in V^{\perp}$. Vamos provar inicialmente dois resultados.

Resultado 4. $V \cap V^{\perp} = \{0\}$.

Prova. *Seja $x = (x_1, ..., x_k) \in V \cap V^{\perp}$. Como x pertence a um subespaço e ao seu complemento ortogonal, deve ser perpendicular a si mesmo. Logo, $< x, x >= x_1^2 + ... + x_k^2 = 0$. Segue então que $x_1 = ... = x_k = 0$.*

Se $\{v_1, ..., v_k\}$ é base de $V \in subset\mathbb{R}^n$, a dimensão de V é igual a k (a dimensão é o número de vetores em uma base). Segue então pelo Resultado 3 que a dimensão de $V^{\perp}$ é igual a $n - k$, ou seja, $V^{\perp}$ tem $n - k$ vetores em uma base. Seja então $\{v_{k+1}, ..., V_n\}$ uma base de $V^{\perp}$. Temos o seguinte resultado:

Resultado 5. *Seja $\{v_1, ..., v_k\}$ uma base de V e $\{v_{k+1}, ..., V_n\}$ uma base de $V^{\perp}$. Então $\{v_1, ..., v_k, v_{k+1}, ..., V_n\}$ é uma base de $\mathbb{R}^n$.*

Prova. *Precisamos provar que o conjunto $\{v_1, ..., v_k, v_{k+1}, ..., V_n\}$ atende a duas condições: é linearmente independente (LI), e gera $\mathbb{R}^n$.*

Vamos provar inicialment que esse conjunto de vetores é LI. Isso significa que precisamos provar que se $c_1v_1 + ... + c_nv_n = 0$, então $c_1 = ... = c_n = 0$. Para tanto, observe que:

$$c_1v_1 + ... + c_kv_k + c_{k+1}v_{k+1} + ... + c_nv_n = 0$$
$$\Rightarrow c_1v_1 + ... + c_k = -(c_{k+1}v_{k+1} + ... + c_nv_n)$$

O lado esquerdo dessa equação é uma combinação linear de vetores em V, e portanto pertence a V (um subespaço é fechado para combinações lineares de seus elementos). Analogamente, o lado direito é uma combinação linear de vetores em $V^{\perp}$, e portanto pertence a $V^{\perp}$. Logo, o vetor $c_1v_1 + ... + c_k = -(c_{k+1}v_{k+1} + ... + c_nv_n)$ pertence à interseção $V \cap V^{\perp}$. Pelo resultado 4, segue que esse vetor é igual à origem. Então:

$$c_1 v_1 + ... + c_k = 0$$

$$-(c_{k+1} v_{k+1} + ... + c_n v_n) = 0$$

Como $\{v_1, ..., v_k\}$ é base de V, segue que $c_1 = ... = c_k = 0$. Analogamente, como $\{v_{k+1}, ..., v_n\}$ é base de $V^\perp$, segue que $c_{k+1} = ... = c_n = 0$. Em suma, concluímos que $c_1 = ... = c_n = 0$, e concluímos que $\{v_1, ..., v_n\}$ é LI.

Como n vetores LI em $\mathbb{R}^n$ geram $\mathbb{R}^n$, concluímos que $\{v_1, ..., v_n\}$ é base para $\mathbb{R}^n$.

Concluímos então que para todo $a \in \mathbb{R}^n$, existe $(c_1, ..., c_n)$ tal que:

$$a = \underbrace{c_1 v_1 + ... + c_k v_k}_{\text{v}} + \underbrace{c_{k+1} v_{k+1} + ... + c_n v_n}_{\text{w}}$$

tal que $v \in V$ e $w \in V^\perp$. Ou seja, todo elemento de $\mathbb{R}^n$ pode ser decomposto como a soma de um elemento em um subespaço V qualquer com um elemento no complemento ortogonal $V^\perp$.

Usaremos a seguinte definição: um conjunto X é a *soma direta* dos conjuntos Y e W se: (i) todo elemento em X por ser representado como a soma de um elemento em Y com um elemento em W; e (ii) a interseção de Y e W é vazia.

Temos então o seguinte corolário:

Resultado 6. *$\mathbb{R}^n$ é igual à soma direta de um subespaço V com seu complemento ortogonal $V^\perp$: $\mathbb{R}^n = V \oplus V^\perp$*

Prova. *Pelo resultado 5, qualquer elemento de $\mathbb{R}^n$ pode ser escrito como a soma de um elemento em V com um elemento em $V^\perp$. Pelo resultado 4, a interseção dos conjuntos V e $V^\perp$ é apenas a origem. Essas duas condições são exatamente a definição de soma direta.*

Resultado 7. *Essa representação é única.*

Prova. *Considere um vetor $a \in \mathbb{R}^n$ tal que $a = v_1 + x_1 = v_2 + x_2$, como $v_i \in V$ e $x_i \in V^\perp$, $i = 1, 2$. Segue então que $v_1 - v_2 = x_1 - x_2$. O lado esquerdo pertence a V (é uma combinação linear de vetores de V), e, analogamente, o lado direito pertence a $V^\perp$. Logo, $v_1 - v_2 = x_1 - x_2 \in V \cap V^\perp = \{0\}$, em que a última igualdade vem do resultado 4. Logo, $v_1 - v_2 = x_1 - x_2 = 0$, e portanto $v_1 = v_2$ e $x_1 = x_2$.*

Podemos provar ainda os seguintes resultados:

Resultado 8. *Seja V um subespaço qualquer. $(V^\perp)^\perp = V$.*

Resultado 9. *Seja $N(A)$ o núcleo associado a uma matriz A.*

$$(N(A))^\perp = ((Col(A^T))^\perp)^\perp = E.C.(A^T = E.L.(A))$$

$$(N(A^T))^\perp = (E.C.(A)^\perp)^\perp = E.C.(A)$$

Solução para um sistema linear no espaço-linha de uma matriz

Considere uma matriz $A_{m,n} = [a_1 a_2 ... a_n]$, em que a_i a i-ésima coluna da matriz A, para $i = 1, ..., n$. Suponha que $b \in E.C.(A)$. Como já visto, existem escalares $x_1, ..., x_n$ tal que $x_1 a_1 + ... + x_n a_n = b$. (Um sistema linear tem solução quando o vetor b de termos livres pertence é combinação linear das colunas da matriz de coeficientes A). Ou seja, existe $x = (x_1, ..., x_n) \in \mathbb{R}^n$ tal que:

$$[a_1 ... a_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = b$$

Esse vetor $x \in \mathbb{R}^n$ é a solução do sistema $Ax = b$. Como já visto, $N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$ e $(N(A))^\perp = E.C.(A^T) = E.L.(A)$. Temos então o seguinte resultado.

Resultado 10. *Para todo b no espaço coluna de A, existe um vetor r_0 no espaço-linha de A que soluciona o sistema linear $Ax = b$. (Ou seja, existe $r_0 \in E.C.(A^T)$ tal que $Ar_0 = b$.)*

Prova. *Uma solução x do sistema linear $Ax = b$ é um vetor em $\mathbb{R}^n$, que pode ser decomposto como a soma direta de um subespaço vetorial qualquer com seu complemento ortogonal. Vamos escolher $\mathbb{R}^n = N(A) \oplus N(A)^\perp = N(A) \oplus E.L.(A) = N(A) \oplus E.C.(A^T)$.*

Logo, qualquer vetor $x \in \mathbb{R}^n$, e em particular a solução do sistema linear $Ax = b$, pode ser escrito como:

$$x = r_0 + n_0$$

Em que $r_0 \in E.L.(A) = E.C.(A^T)$ e $n_0 \in N(A)$. Podemos então escrever $r_0 = x - n_0$. Pré-multiplique ambos os lados para obter:

$$Ar_0 = A(x - n_0) = Ax - An_0$$

Como n_0 está no espaço-nulo de A, segue que $An_0 = 0$, e portanto $Ar_0 = Ax = b$, em que a última igualdade vale porque estamos trabalhando com um vetor x que soluciona o sistema $Ax = b$. Em suma, $Ar_0 = b$, e portanto $r_0 \in E.C.(A^T)$ soluciona o sistema linear $Ax = b$.

É possível provar ainda que essa solução r_0 é única no espaço-linha de A.

Resultado 11. *r_0 é único.*

Prova. *Suponha que existam $r_0, r_1 \in E.C.(A^T)$ tal que $Ar_0 = Ar_1 = b$. Então $r_1 - r_0 \in E.C.(A^T)$, pois o espaço-coluna é um subespaço vetorial e portanto é fechado para combinações lineares; e além disso $A(r_1 - r_0) = b - b = 0$, e portanto $(r_1 - r_0) \in N(A) = (E.C.(A^T)^\perp$. Logo, $_1 - r_0 \in E.C.(A^T) \cap (E.C.(A^T))^\perp$. Pelo resultado 4, segue que $r_1 - r_0 = 0$, e portanto $r_1 = r_0$.*

Dessa forma, qualquer solução x do sistema linear $Ax = b$ pode ser escrita como uma soma $r_0 + n_0$, em que r_0 é fixo e n_0 varia. É um resultado já visto: as soluções de um sistema linear não-homogêneo podem ser escritas como a soma de uma solução particular r_0 como uma solução n_0 do sistema linear homogêneo associado.

A solução r_0 do sistema linear $Ax = b$ tem uma propriedade relevante: ela é a mais próxima da origem dentre todos os vetores x que solucionam $Ax = b$.

Resultado 12. *Sejam x e r_0 duas soluções do sistema linear $Ax = b$, e $r_0 \in E.C.(A^T)$. Então $||x|| \geq ||r_0||$.*

Prova. *Como visto, todo x que soluciona $Ax = b$ pode ser escrito como $r_0 + n_0$. Então:*

$$||x||^2 =< x, x >=< r_0 + n_0, r_0 + n_0 >=< r_0, r_0 > + \underbrace{2 < r_0, n_0 >}_{0} + < n_0, n_0 >$$

$$= ||r_0^2|| + ||r_0^2||$$

Usamos $< r_0, n_0 >= 0$ pois r_0 e n_0 são ortogonais. Para concluir, observamos que $||r_0^2||$ é não negativo, e portanto $||x||^2 \geq ||r_0||^2$. Logo, $||x|| \geq ||r_0||$.

Projeção em Subespaços

Como visto, para todo vetor $x \in \mathbb{R}^n$, $\exists!(v, w) \in V \times V^{\perp}$ tal que $x = v + w$. (Lembrando: o símbolo $\exists!$ significa "existe um único").

Definimos então a projeção ortogonal de x no subespaço V como:

$$\underset{V}{Proj}(x) = v$$

Analogamente, definimos:

$$\underset{V^{\perp}}{Proj}(x) = w$$

Podemos encontrar uma expressão para $\underset{V}{Proj}(x) = v \in V$. Seja $\{b_1, ..., b_k\}$ uma base para V. Logo, para todo $a \in V$, existe $(y_1, ..., y_k)$ tal que:

$$a = y_1 b_1 + ... + y_k b_k$$

(Todo elemento do subespaço pode ser escrito como combinação linear de uma base desse subespaço.) Forme então uma matriz $A_{n,k}$ cujas colunas são os vetores da base acima:

$$A_{n,k} = \begin{bmatrix} | & ... & | \\ b_1 & ... & b_k \\ | & ... & | \end{bmatrix}$$

Resultado 13. *Seja A uma matriz cujas colunas são os vetores de uma base do subespaço vetorial V. Seja Ay a projeção de $x \in \mathbb{R}^n$ no subespaço V. O vetor y é dado pela seguinte expressão:*

$$y = (A^T A)^{(-1)} A^T x$$

Prova. *Por definição, $V = E.C.(A)$. Podemos então escrever $a = Ay = b_1 y_1 + ... + b_k y_k$: qualquer elemento $a \in V$ pode ser escrito como Ay, ou seja, como combinação linear dos vetores da base de A.*

Mas vimos acima que $\underset{V}{Proj}(x) \in V$. Logo, $\underset{V}{Proj}(x)$ também pode ser escrito dessa forma: para algum $y \in \mathbb{R}^n$, é verdade que:

$$\underset{V}{Proj}(x) = Ay$$

Também vimos que podemos escrever $x \in \mathbb{R}^n$ *como uma soma:* $x = \underset{V}{Proj}(x) + w$*, em que* w *é um vetor (único) de* $V^{\perp}$*. Então:*

$$x - \underset{V}{Proj}(x) = w$$

Como $w \in V^{\perp}$*, segue que* $x - \underset{V}{Proj}(x) \in V^{\perp}$*. Sabemos também que* $V = E.C.(A)$*, e portanto* $V^{\perp} = E.C.(A)^{\perp} = N(A^T)$*, e portanto* $x - \underset{V}{Proj}(x) \in V^{\perp} = N(A^T)$*. Então:*

$$A^T(x - \underset{V}{Proj}(x)) = 0$$

$$A^T x - A^T \underbrace{\left(\underset{V}{Proj}(x)\right)}_{Ay} = 0$$

$$A^T x - A^T A y = 0$$

$$A^T x = (A^T A) y$$

Como as colunas de A *são linearmente independentes, é possível provar que* $(A^T A)$ *admite inversa. Então:*

$$(A^T A)^{(-1)} A^T x = y$$

Podemos então escrever:

$$\underset{V}{Proj}(x) = Ay = A(A^T A)^{(-1)} A^T x$$

Essa é a expressão da projeção de $x \in \mathbb{R}^n$ *no subespaço vetorial* V*. Como* $A(A^T A)^{(-1)} A^T$ *é apenas um produto de várias matrizes, essa expressão representa uma transformação linear - ou seja, a projeção ortogonal é uma transformação linear.*