autovalores e auto vetores

Jacques Wainer

18/3/20

Um outro livro texto:
https://www.ufrgs.br/reamat/AlgebraLinear/livro/main.html

Mudança de base

$x = [1, 2, 3]$ é o ponto 1*[1,0,0]+*2[0,1,0]+3*[0,0,1]

aula seriam as coordenadas deste *mesmo ponto* se a base fosse $b1 = [1, 4, 2]$ $b2 = [0, 3, 7]$ e e $b3 = [9, 0, -5]$?

Se y fossem essas coordenadas e se

$$P = [b1 \quad | \quad b2 \quad | \quad b3] = \begin{bmatrix} 1 & 0 & 9 \\ 4 & 3 & 0 \\ 2 & 7 & -5 \end{bmatrix}$$

P representa as novas bases como uma matriz. as novas bases são as coluna de P

então

$$x = Py$$

ou seja

$$y = P^{-1}x$$

P precisa ser fullrank (não singular) para ter inversa

Mudança de base para uma matriz

Se a matrix A faz uma certa transformação $y = Ax$

Como seria uma matriz que faz a mesma transformação que o A mas na base nova?

Como obter um B tal que $y' = Bx'$ onde x' é o x na nova base, e y' é o y na nova base.

$B = P^{-1}AP$

então:

$$\begin{gathered} y = Ax \\ P^{-1}y = P^{-1}Ax \\ P^{-1}y = P^{-1}AIx \\ P^{-1}y = P^{-1}A(PP^{-1})x \\ P^{-1}y = P^{-1}AP(P^{-1}x) \\ y' = P^{-1}APx' \end{gathered}$$

Se P é a matriz das nova base,

$$B = P^{-1}AP$$

Matrizes ortogonais

se a mudança de base é para uma base ortonormal (bases de tamanho 1 e ortogonais entre si

A matriz P é chamada de ortogonal (as colunas são ortonormais) e é usualmente representada por Q

$$Q^{-1} = Q^T$$

Transformadas por matrizes ortogonais são rotações e talvez reflexões (uma coodenada x_i passa a ser $-x_i$)

Autovetores e autovalores

Numa transformada linear A pode acontecer que

$$Ax = \lambda x$$

ou seja a transformação apenas estica ou comprime o vetor x.

neste caso

- x é um autovetor (eigenvector)
- α é o autovalor (eigenvalue) associado ao autovetor.

Se x é um autovetor então $2x$ também é

$A2x = 2(Ax) = 2\lambda x = \lambda(2x)$

Chamamos de autovetor apenas vetores com tamanho = 1 ($|x| = 1$).

Equação característica

Para x um autovetor

$$Ax = \lambda x$$
$$Ax - \lambda x = 0$$
$$Ax - \lambda Ix = 0$$
$$(A - \lambda I)x = 0$$

Considere que pode haver mais de um autovetor/autovalor. Todos eles satisfazem a equação, onde λ é a "variável".

Nos não discutimos determinante neste curso, mas a equação implica

$$\det(A - \lambda I) = 0$$

Se A é nxn então a equação característica é de grau n, e portanto pode ter até n soluções.

Uma vez que obtém-se uma solução λ_i pode-se computar o autovetor x_i associado usando-se a equação $(A - \lambda_i I)x_i = 0$ ou seja x_i esta no nullspace da matriz $(A - \lambda_i I)$

Ex:

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

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

$$det(A - \lambda I) = (1-\lambda)^2 - 4 = 1 - 2\lambda + \lambda^2 - 4 = \lambda^2 - 2\lambda - 3 = 0$$

$$\lambda = 3 \quad e \quad \lambda = -1$$

Multiplicidade de autovalor

Resolver a equação característica pode gerar uma mesma solução mais de uma vez. Por exemplo quando uma equação quadrada só tem uma solução ($\Delta = 0$).

Isso é chamado de multiplicidade do autovalor

Se um autovalor tem multiplicidade > 1 então ele tem pelo menos 2 autovetores linearmente independentes associados a ele.

A dimensão do nullspace associado a um autovalor (o nullspace de $A - \lambda_i I$) é chamado de multiplicidade geometrica do autovalor

De vez em quando a multiplicidade geometrica é menor que a multiplicidade do autovalor

Ordem dos autovetores

Convencionalmente ordena-se os autovetores por ordem decrescente de modulo do autovalor.

Assim o autovetor associado ao autovalor de maior modulo é o **primeiro** autovetor, o vetor associado ao 2o maior autovalor em modulo é o 2o autovetor, e assim por diante.

Eigengap

eigengap é a diferença do maior para o 2o maior autovalor.

Se o eigegap é grande, então há um procedimento simples para calcular o primeiro autovetor (associado ao maior autovalor)

1. crie um vetor aleatorio x_1 de tamanho 1
2. $x_2 = Ax_1$,
3. normalize x_2 para ter tamanho 1
4. repita o procedimento poucas vezes

O componente de x_1 na direção do primeiro autovetor vai ser esticado muito mais que as outras direções (que no maximo serão esticadas por λ_2)

Em poucas iterações o vetor x_i se aproxima do primeiro autovetor.

Diagonalização

Em alguns casos é possível fatorar a matriz A como

$$A = PDP^{-1}$$

onde D é uma matriz diagonal dos autovalores a matriz A

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}$$

P é a matriz de mudança de base que leva as bases para os autovetores de A

$P = [v_1 \quad | \quad v_2 \quad | \quad v_3 \quad | \quad ...v_n]$

Condições:

- A possui n autovalores reais distintos
- A para os autovalores com multiplicidade > 1 a multiplicidade geometrica é igual a multiplicidade.

se $A = PDP^{-1}$ então

$$\begin{aligned} A^{-1} = (PDP^{-1})^{-1} &= (P^{-1})^{-1} D^{-1} P^{-1} \\ &= PD^{-1} P^{-1} \end{aligned}$$

Onde D^{-1} é uma matriz diagonal com elementos $\frac{1}{\lambda_i}$

Matrizes simétricas

no caso de matrizes simétricas

- são diagonalizáveis
- os autovetores são ortogonais entre si
- a matrix P é ortogonal e portanto
- $A = P^T DP$

Autovetores em grafos

Matrizes quadradas são uma representação de grafos. Mas não é claro o que é um vetor nessa situação.

[1,0,0,0] é o vertice 1, mas o que é [2,5,-3,0]?

Uma interpretação é que os números são "cargas" que estão em cada vertice, e multiplicar o grafo e o vetor é transferir as cargas através das arestas.

Se as arestas tem peso, então as cargas são multiplicadas pelos pesos

Um autovetor então é uma distribuição de cargas que se repete (com a multiplicação de λ_i) quando as cargas são transferidas pelo grafo.

Alguns padrões

$x^T y$

- $x^T y$ é o produto escalar de x e y

$x^T Ax$

- $x^T Ax$ e A é simétrica é algo como os termos quadráticos de uma equação

Se x tem 3 dimensões, então so termos quadráticos sao

$a_0 x_1^2 + a_1 x_2^2 + a3x_3^2 + a_4 x_1 x_2 + a_5 x_1 x_3 + a_6 x_2 x_3$

onde os a_i vem da matrix A (simétrica)

Da aula passada, uma distribuição normal de uma variável

$$N(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2}\frac{x^2-\mu^2}{\sigma^2}}$$

Para um x que é um vetor

$$N(x|\mu, \Sigma) = Ke^{\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

https://en.wikipedia.org/wiki/Multivariate_normal_distribution

$P^{-1}BP$ e $P^T BP$

- $P^{-1}BP$ é uma mudança de base para uma transformada de tal forma que a matriz B é mais conveniente ou simples nesta base nova

- $P^T BP$ é uma mudança de base onde a mudança é uma rotação com possivelmente reflexões

$y^T Ax$

- eu interpreto essa expressão como sendo uma deformação da medida de tamanho (ou distancia)

- se A = I entao é o produto escalar tradicional, que é uma medida de similaridade e ou distancia

- se A é diagonal entao é um produto escalar com coeficientes em cada coordenada que dao a importancia da coordenada.
- se A é uma matriz em geral, então eu dou esse nome de deformação