Curso: Aprendizagem de Máquina em Inteligência Artificial

Disciplina: Aprendizado Não Supervisionado

Prof. Marcelo Novaes de Rezende

PCA

Principal Component Analysis (PCA)

Faculdade de Computação e Informática
Mackenzie

Antes, vamos ver a base matemática de PCA

Autovalores e Autovetores

Autovalores e Autovetores

Seja M uma matriz, v um vetor e λ um escalar.

Dizemos que v é um autovetor de M e λ é seu autovalor associado quando:

$$Mv = \lambda v$$

Assim, a transformação linear causada por M em um vetor é equivalente ao produto de um escalar por esse vetor...vamos destacar alguns pontos:

1)o produto de um escalar λ (>0) por um vetor gera outro vetor paralelo ao inicial. Por exemplo, o vetor (x,y)=0,4 está em y. Se o multiplicarmos por 2, será (0,8), ainda no eixo y.

2)Devemos sempre pensar em pares (autovalor, autovetor)

3)A matriz M é quadrada

4)Autovalores podem ser números reais ou imaginários

Faculdade de Computação e Informática
Mackenzie

Autovalores e Autovetores

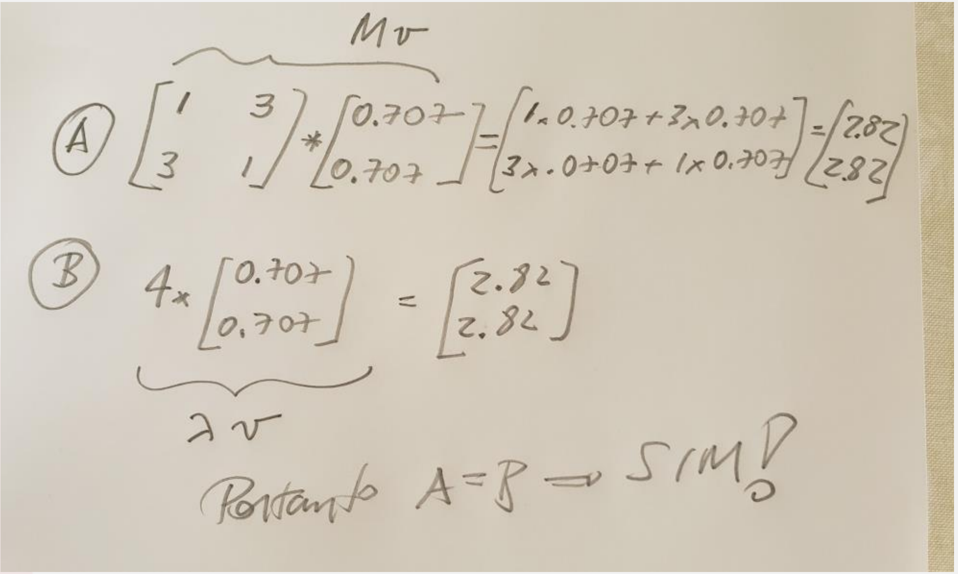
Pergunta:

O vetor v=(0.707,0.707) é um autovetor da matriz

$$M=\begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} \quad ?$$

4 é o autovalor associado?

Em Python, vamos ver **pca1.ipynb**



Faculdade de Computação e Informática
Mackenzie

Autovalores e Autovetores

O vetor v=(0.707,0.707) é um autovetor da matriz

$$M=\begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix}$$

Vamos aplicar a transformação da matriz a pontos de um “grid” 10x10..e observar o resultado...em Python.

Faculdade de Computação e Informática
Mackenzie

Autovalores e Autovetores...como obtê-los?

$$Mv = \lambda v$$

$$(M - \lambda I)v = 0 \quad (1)$$

Obs: I é a matriz identidade..

Obs: 0 é um vetor de "m" zeros.

Obs: λ é um escalar (número)

Obs: M é uma matriz quadrada mxm

Obs: $M - \lambda I$ é uma matriz quadrada mxm

Obs: v é um vetor com m componentes

Essa equação **1** (sistema de equações lineares homogêneo) tem uma solução trivial por ser um produto com resultado um vetor zero (vetor v=0). Para que existam outras soluções além da trivial, o determinante de $M - \lambda I$ deve ser nulo. Dessa imposição, chegamos em uma equação que nos permite encontrar os autovalores e, depois, os autovetores (obviamente).

Autovalores e Autovetores

$$M = \begin{bmatrix} 1 & -2 \\ 3 & 9 \end{bmatrix}, \quad Mv = \lambda v, \quad \underbrace{(M-\lambda I)v = \emptyset}_{I}$$

Impor $\det[M-\lambda I] = \emptyset$ para que a equação I tenha mais que uma solução:

$$M - \lambda I = \begin{bmatrix} 1 & -2 \\ 3 & 9 \end{bmatrix} - \lambda \begin{bmatrix} 1 & \emptyset \\ \emptyset & 1 \end{bmatrix} = \begin{bmatrix} 1-\lambda & -2 \\ 3 & 9-\lambda \end{bmatrix}$$

$$\det \begin{bmatrix} 1-\lambda & -2 \\ 3 & 9-\lambda \end{bmatrix} = (1-\lambda)(9-\lambda) - (-2 \times 3)$$

$$\det = 9 - \lambda - 9\lambda + \lambda^2 + 6 = 0$$

$$\lambda^2 - 10\lambda + 15 = 0 \rightarrow \lambda_1 \cong 8.16$$

$$\lambda_2 \cong 1.84$$

Solução da equação de 2º grau!

Exemplo: Obter os **autovalores** de

$$M = \begin{vmatrix} 1 & -2 \\ 3 & 9 \end{vmatrix}$$

No Python, **pca2.ipynb**

Faculdade de Computação e Informática
Mackenzie

Autovalores e Autovetores

Para λ=8.16, o autovetor correspondente é :

(M- λI)v=0

$$\begin{matrix} -7.16 & -2 \\ 3 & 0.84 \end{matrix} \quad \begin{matrix} x \\ y \end{matrix} = \begin{matrix} 0 \\ 0 \end{matrix}$$

-7.16x-2y=0

Exemplo: para x=1 v={ 1, -3.58} Obs: há infinitos autovetores paralelos ao primeiro....

Teste: $\begin{matrix} 1 & -2 \\ 3 & 9 \end{matrix} \quad \begin{matrix} 1 \\ -3.58 \end{matrix} = 8.16 \begin{matrix} 1 \\ -3.58 \end{matrix}$?   Vamos ao Python **(pca2.ipnynb)**

Autovalores e Autovetores

Analogamente, para λ=1.84, o autovetor correspondente é :

(M- λI)v=0

$$\begin{matrix} -0.84 & -2 \\ 3 & 7.16 \end{matrix} \quad \begin{matrix} x \\ y \end{matrix} = \begin{matrix} 0 \\ 0 \end{matrix}$$

-0.84x– 2y =0
y=-0.42x

Exemplo: para x=1 v={ 1, -0.42}

Teste: $\begin{matrix} 1 & -2 \\ 3 & 9 \end{matrix} \quad \begin{matrix} 1 \\ -0.42 \end{matrix}$ = 1.84 $\begin{matrix} 1 \\ -0.42 \end{matrix}$? Sim! Vamos ao Python **(pca2.ipnynb)**

Autovalores e Autovetores

É interessante deixar os autovetores com módulo 1. Para isso, basta obter a norma e dividir seus componentes por essa norma...

λ=8.16 o vetor v={1, -3.58} tem norma: 3.72
e fica v={ 0.27, -0.96) com norma 1

λ=1.84 o vetor v={1, -0.42} tem norma: 1.08
e fica v={ 0.92, -0.39) com norma 1

Autovalores e Autovetores

Propriedade Importante:

O produto dos autovalores é o determinante da matriz:

M= $\begin{matrix} 1 & -2 \\ 3 & 9 \end{matrix}$ det(M)= 9+6=15 = 8.16*1.84

Vamos testar tudo isso no Python! **pca3.ipynb**

Faculdade de Computação e Informática
Mackenzie

Correlação Pearson e Spearman
Vamos analisar os notebooks

Correlations*.ipynb

Hermitian Matrix

A matriz M é 'Hermitiana" se a transposta da conjugada é a própria M (?)

Exemplos:

$$\begin{bmatrix} 3 & 2+i & -i \\ 2-i & 5 & 5-i \\ i & 5+i & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 2 & -i & 10+i \\ 2 & 1 & 0 & -5i \\ i & 0 & -2 & 0 \\ 10-i & 5i & 0 & 5 \end{bmatrix} \quad \begin{bmatrix} 0 & i & i \\ -i & 0 & i \\ -i & -i & 0 \end{bmatrix}$$

É fácil identificá-las: Na diagonal principal só há reais e os elementos simétricos são conjugados (a+bi , a-bi)

Toda matriz real simétrica é Hermitiana.

Toda Matriz hermitiana tem autovalores reais

Os autovetores associados a autovalores diferentes são ortogonais

Matriz de Covariância

Covariância:

$$\sigma(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

Matriz de covariância
(para 3 features)

$$\begin{bmatrix} Var_1 & Cov_{1,2} & Cov_{1,3} \\ Cov_{1,2} & Var_2 & Cov_{2,3} \\ Cov_{1,3} & Cov_{2,3} & Var_3 \end{bmatrix}$$

Matrizes de Covariância são reais e simétricas, portanto Hermitianas

Todas as matrizes de covariância são positivas semi definidas. Isso implica que todos os autovalores são positivos.

Uma matriz positiva semi definida A => $v_t Av>=0$ para qualquer v diferente de zero

Vamos ver em Python pca4.ipynb

PCA : Principal Component Analysis

A Análise de Componentes Principais (PCA) é um
ferramenta de redução de dimensionalidade que pode ser
usada para
reduzir um grande conjunto de variáveis para um pequeno
conjunto
que ainda contém a maioria das informações do
conjunto grande

PCA : Principal Component Analysis

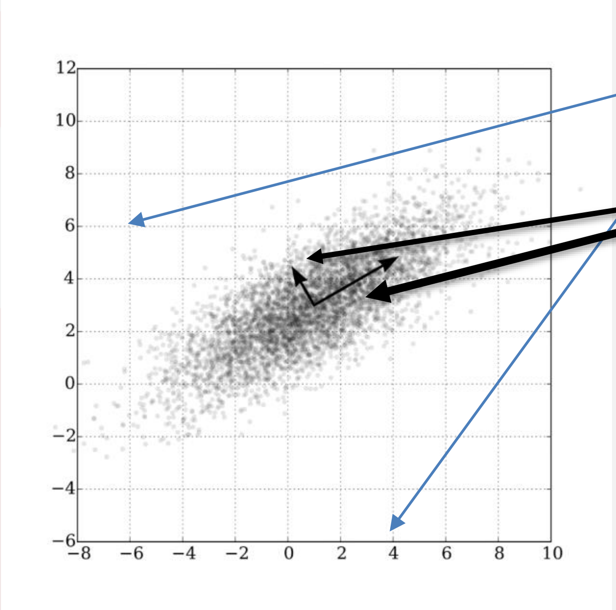
Basicamente, partimos da matriz de covariância entre as features e extraímos seus autovalores e autovetores correspondentes.

Os autovetores associados a maiores autovalores indicarão maior "variância naquele eixo" => mais importância...mais variância, mais informação

Podemos representar o dataset nesses eixos dos autovetores com menos dimensões que o inicial desprezando os eixos com autovalores baixos.

É possível voltar ao dataset incial, mas com perdas....

PCA : Principal Component Analysis



Eixos originais

Eixos de maior variância

Faculdade de Computação e Informática
Mackenzie

PCA : Principal Component Analysis

Partindo de pca5.ipynb

1)Analisar o código com scikit

2)Verificar a matriz de covariância antes e depois das transformações

3)Calcular a acurácia (logistic regression) para classificação multiclasse (3) só com dois componentes (após PCA) e com 4 componentes (Original)

4)Reconstruir o dataset (de 2 para 4 features) e observar perdas (é uma lossy compression)

5)Fazer a visualização 2D das 3 classes após PCA

Usar **pca5.ipynb**

PCA: Partindo de pca6.ipynb, e sem standardizar as features, veja o mínimo de features necessárias (Obtidas por PCA) para se ter mais de 80% de acurácia com regressão logística na classificação de fraude ou não.

Até a próxima aula

OBRIGADO!

Prof Marcelo Rezende
email rezendemn@gmail.com