### SCC0270/SCC5809 - Redes Neurais Aula 12 - PCA e Rede PCA

Profa. Dra. Roseli Aparecida Francelin Romero SCC - ICMC - USP

2017

#### Sumário

- PCA clássica
- Caracterização
- ExemploRedução de dimensionalidade
- Algoritmo
- 2) Rede PCA adaptativa

- Dado um conjunto de amostras, cada qual com um conjunto finito de variáveis.
- Derivar novas componentes que produzam uma descrição mais simples do sistema.
- Reduzir as variáveis originais a um número menor de variáveis ortogonais (não correlacionadas).
- Mudança de espaço de variáveis.

#### Objetivo

Dadas p variáveis, deseja-se achar combinações lineares dessas para produzir índices que **não** sejam correlacionados, de tal forma que:

• Indices Z: componentes principais.

PCA clássica

#### Sumário

- PCA clássica
- Caracterização
- Exemplo
- Redução de dimensionalidade
- Algoritmo
- 2 Rede PCA adaptativa

i-ésima componente principal.

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p$$

Com restrição:

$$a_{i1}^2 + a_{i2}^2 + \dots + a_{ip}^2 = 1$$

• E com  $Z_1, Z_2, \ldots, Z_{i-1}, Z_i$  não correlacionados.

• PCA: resume-se em encontrar os autovalores e autovetores da matriz C de covariância dos dados.

$$= \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{bmatrix}$$

 Supondo que os autovalores da matriz C estejam ordenados da seguinte forma:

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_j \geq \cdots \geq \lambda_p$$

Os autovetores associados são:

$$a_1, a_2, \cdots, a_j, \cdots, a_p$$

Propriedades:

$$a_i^{\mathcal{T}}a_j=egin{cases} 1, & i=j \ 0, & i
eq j \end{cases}$$

Para:

$$Z_i = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p$$

•  $(a_{i1}, a_{i2}, \ldots, a_{ip})$  são os elementos do i-ésimo autovetor correspondente.

 A soma dos autovalores corresponde ao traço da matriz covariância C:

$$\lambda_1 + \lambda_2 + \dots + \lambda_p = c_{11} + c_{22} + \dots + c_{pp}$$

•  $var(Z_1) \geq var(Z_2) \geq \cdots \geq var(Z_p)$ 

$$var(Z_i) = \lambda_i$$

#### Sumário

- PCA clássica
- Caracterização
- Exemplo
- Redução de dimensionalidade
- Algoritmo
- 2) Rede PCA adaptativa

11/45

## Exemplo: conjunto Iris.dat

Tabela 1: Aplicando PCA na base de dados Iris.dat

			Autovetores	Autovetores (coeficientes)	
Componente Autovalor	Autovalor	$\chi_{ m I}$	$X_2$	$X_3$	$X_4$
1	2.91082	0.522371	-0.263356	0.581254	0.565611
2	0.92122	0.372320	0.925556	0.021094	0.065417
3	0.14735	-0.721015	0.242033	0.140889	0.633804
4	0.02061	-0.261998	0.124137	0.801155	-0.523543

## Exemplo: conjunto Iris.dat

#### Conclusões:

- ullet Z $_1$  é responsável por 72.77% do total da variância.
- $Z_2$  é responsável por 23.03% do total da variância.
- Z<sub>3</sub> é responsável por 3.68% do total da variância.
  - $Z_4$  é responsável por 0.52% do total da variância.

#### Exemplo

## Reconstrução dos dados originais

$$Z = [Z_1, Z_2, \dots, Z_p]^T$$
  
=  $[X^T a_1, X^T a_2, \dots, X^T a_{p-1}]^T$   
=  $A^T X$ 

matriz ortogonal 
$$\rightarrow A^T = A^{-1}$$

$$X = A \cdot Z = \sum_{i=1}^{p} Z_i a_i$$

PCA clássica

#### Sumário

- PCA clássica
- Caracterização
- Exemplo
- Redução de dimensionalidade
- Algoritmo
- 2 Rede PCA adaptativa

15/45

PCA clássica

## Redução de dimensionalidade

• Sejam  $\lambda_1, \lambda_2, \ldots, \lambda_m$  os m autovalores da matriz C.

• Então,  $X' \sim X$ , onde:

$$X' = \sum_{i=1}^{m} Z_i a_i \qquad m < \mu$$

• Erro: e = X - X', de modo que:

$$e = \sum_{i=m+1}^{p} Z_i a_i$$

• O vetor de erro e é ortogonal ao vetor X', que aproxima X.

• Princípio da ortogonalidade:  $e^T X' = 0$ 

#### Aplicação

Tabela 2: Taxas de compressão para blocos 8x8.

Número de	Dimensão do Bloco	Taxa de Compressão
Componentes	Compactado	
Principais		
1	8x1	$1 - \left(\frac{1}{8}\right) = 87,5\%$
2	8x2	1 - (2/8) = 75%
3	8x3	$1 - \left(\frac{3}{8}\right) = 62,5\%$
4	8x4	$1 - \left(\frac{4}{8}\right) = 50\%$

18/45

#### Aplicação

Tabela 3: Taxas de compressão para blocos 16x16.

Taxa de Compressão			$1 - (\frac{1}{16}) = 93,75$	1 - (2/16) = 87,59
Dimensão do Bloco	Compactado		16x1	16x2
Número de	Componentes	Principais		2

#### Sumário

- PCA clássica
- Caracterização
- ExemploRedução de dimensionalidade
- Algoritmo
- 2) Rede PCA adaptativa

20/45

#### Algoritmo

- Padronize os dados correspondentes às variáveis para que estes tenham média igual a 0 e variância igual a 1.
- Calcule a matriz de correlação C.
- autovetores. Os coeficientes da i-ésima componente principal Encontre os autovalores da matriz C e seus correspondentes são dados pelo autovetor associado ao i-ésimo autovalor.
- Descarte as componentes que acumulem uma pequena proporção da variação dos dados.
- 90% do total da variância, as outras 17 componentes principais três primeiras componentes principais forem responsáveis por • Por exemplo, se os dados originais tiverem 20 variáveis e as podem ser ignoradas.

# Normalizando os dados (passo 1 do algoritmo)

Dados centrados na média e variância 1:

$$ar{x} = rac{1}{n} \sum_{i=1}^n x_i$$

$$x_i \leftarrow \frac{(x_i - \bar{x})}{var(x_i)}$$

 Escalonamento pela variância (quando há variáveis muito dominantes em relação às demais):

$$var(x_j) = \frac{1}{n-1} \sum_{i=1}^{n} (x_{ji} - \bar{x_j})^2$$
$$x_i \leftarrow \frac{x_i}{var(x_i)}$$

#### Sumário

- 1) PCA clássica
- Rede PCA adaptativa
- RNA para compressão de imagens
  - Regra de Hebb
- Convergência
- Algoritmo
- Experimentos e resultados
- Conclusões

0000000000000000000

PCA clássica

RNA para compressão de imagens

#### Sumário

- 1) PCA clássica
- 2 Rede PCA adaptativa
- RNA para compressão de imagens
  - Regra de Hebb
- Convergência
- Algoritmo
- Experimentos e resultados
- Conclusões

24/45

## RNA para compressão de imagens

- Padrão JPEG: mais utilizado.
- PCA Clássica: método estatístico multivariado.
- Rede PCA Adaptativa: arquitetura de Redes Neurais Artificiais.

#### Motivação

• Redução da quantidade de dados armazenadas em sistemas computacionais. • Redução da dimensionalidade de imagens que ocupem grande quantidade de memória.

Obtenção de métodos que:

Atinjam altas taxas de compressão.

Não prejudiquem a qualidade visual.

PCA clássica ooooooooooooooo

#### Sumário

- 1) PCA clássica
- 2 Rede PCA adaptativa
- RNA para compressão de imagens
- Regra de Hebb
- Convergência
- Algoritmo
- Experimentos e resultados
- Conclusões

27/45

#### Regra de Hebb

#### Postulado de Hebb

processo crescente ou mudanças metabólicas ocorrem em ambas as "Quando um axônio da célula A está suficientemente próximo para excitar uma célula B e repetidamente tenta excitá-la, algum

- Transformando em regras:
- Se dois neurônios ligados por uma sinapse são simultaneamente ativados, a intensidade dessa sinapse(conexão) é aumentada.
- assincronamente, a intensidade dessa sinapse é diminuída ou Se dois neurônios ligados por uma sinapse são ativados até mesmo eliminada.

#### Regra de Hebb

Regra de Hebb:

$$\Delta w_{kj}(n) = \eta y_k(n) x_j(n)$$

Onde:

•  $\eta$  é uma constante positiva  $\rightarrow$  velocidade de aprendizado.

•  $y_k(n)$  é a saída do neurônio k no tempo n.

•  $x_j$  é o j-ésimo elemento do vetor de entrada no tempo n.

Regra anti-hebbiana:

$$\Delta w_{kj}(n) = -\eta y_k(n) x_j(n)$$

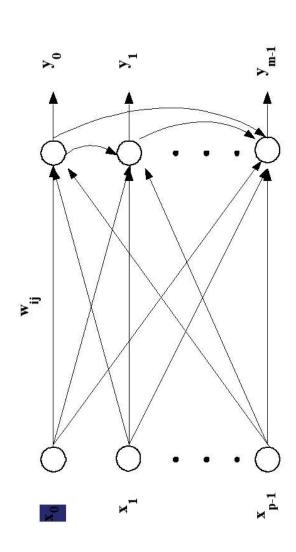
#### Sumário

- 1 PCA clássica
- Rede PCA adaptativa
- RNA para compressão de imagens
  - Regra de Hebb
- Convergência
- Algoritmo
- Experimentos e resultados
- Conclusões

Rede PCA adaptativa

Convergência

### Rede PCA adaptativa



Saída: 
$$y_j(n) = \sum_{i=0}^{p-1} w_{ij}(n)x_i(n) + \sum_{l < j} u_{lj}(n)y_l(n),$$

Ajuste dos pesos: 
$$\Delta w_{ij}(n) = \eta \kappa_i(n) y_j(n)$$

Ajuste dos pesos laterais: 
$$\Delta u_{lj}(n) = -\mu y_l(n) y_j(n)$$

Convergência

## Teorema de convergência

#### eorema

generalizada de Hebb irá convergir na média e, no limite, irá se aleatórios no tempo n=0, então, com probabilidade 1, a regra autovetores da matriz C de covariância dos vetores de entrada Se a matriz de pesos sinápticos W(n) for associada a valores aproximar de uma matriz cujas colunas serão os primeiros m x(n), ordenados por ordem decrescente de autovalor.

## leorema de convergência

Portanto, no limite, pode-se escrever:

$$\Delta w_j(n) \to 0, \quad w_j \to a_j \qquad j = 0, 1, \dots, m - 1$$

- Tal que  $||w_j(n)|| = 1$  para todo j.
- covariância dos vetores de entrada x(n), estando esses associados aos *m* maiores autovalores da matriz C de Os valores representam os autovetores normalizados autovalores ordenados em ordem decrescente.

### leorema de convergência

 Pode-se acelerar a convergência da rede introduzindo um aprendizagem e o momentum diminuam com o tempo. termo momentum  $\beta$  e deixando que os parâmetros de

$$\Delta w_{ij}(n+1) = \eta(n)x_iy_i + \beta(n)\Delta w_{ij}(n)$$
 (1)

$$\Delta u_{ij}(n+1) = -\mu(n)y_iy_j + \beta(n)\Delta u_{ij}(n) \tag{2}$$

• 
$$\eta(n+1) = \max\{\alpha\eta(n), 0.0001\}$$
  
•  $\mu(n+1) = \max\{\alpha\eta(n), 0.0002\}$ 

$$\mu(n+1) = \max\{\alpha \eta(n), 0.0002\}$$

• 
$$\beta(n+1) = \max\{\alpha\eta(n), 0.0001\}$$

• 
$$\alpha$$
 é o fator de limitação.

#### Sumário

- 1 PCA clássica
- 2 Rede PCA adaptativa
- RNA para compressão de imagens
  - Regra de Hebb
- Convergência
- Algoritmo
- Experimentos e resultados
- Conclusões

#### Algoritmo

- Inicialize todos os pesos de conexões com pequenos valores aleatórios, escolha os valores para os parâmetros de aprendizagem e normalize-os em [0, 1].
- Se normalizar em [-1,1], pode mudar os sinais dos autovetores.
- Repita:
- Selecione aleatoriamente um padrão p-dimensional e apresente-o à rede.
- $\sigma$ Ajuste os pesos das conexões entre a camada de entrada e camada de saída, de acordo com a equação 1.
- Normalize os vetores-peso (em colunas).
- Atualize os pesos laterais, de acordo com a equação 2 (não precisa normalizar).
- **6** Modifique os parâmetros  $\beta$ ,  $\eta$  e  $\mu$ .

pequenos OU um número de iterações máximo seja atingido. Até que todos os pesos laterais sejam suficientemente

0000000000000000000

PCA clássica

#### Sumário

- 1) PCA clássica
- 2 Rede PCA adaptativa
- RNA para compressão de imagens
  - Regra de Hebb
- Convergência
- Algoritmo
- Experimentos e resultados
- Conclusões

#### Experimentos

- Dados sobre as imagens:
- Conjunto composto por 208 imagens médicas.
- Dimensão: 480x640 pixels (valor em nível de cinza).
- Representação de cortes de um fígado humano
- aquisição a partir de um microscópio laser, pelo Departamento de Patologia da Fiocruz.

# Compressão de imagens através do JPEG

Utilização do aplicativo xv (xview), para Linux.

Possui módulo para compressão JPEG.

Permite que o usuário selecione taxa de compressão.

### Taxas de compressão

comprimento da cadeia de dados comprimidos  $\mathit{Taxa}$  de Compress $ilde{a}o = 1$  – –

comprimento da cadeia de dados originais

## Taxas de Compressão para Blocos 32x32

		6	
Taxa de Compressão	$1 - \left(\frac{1}{32}\right) = 96,875\%$	1 - (2/32) = 93,75%	$1 - \left(\frac{3}{32}\right) = 90,625\%$
Número de Componentes Principais	I	2	3

Experimentos e resultados

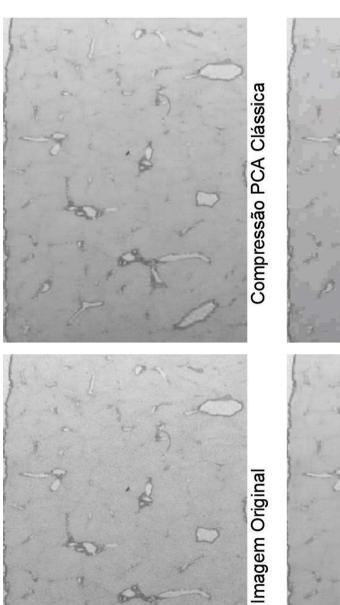
#### Resultados

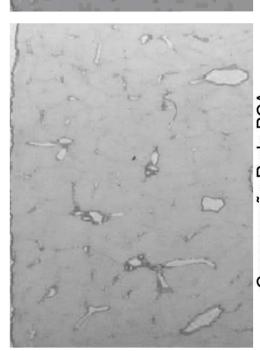
				Taxa d	Taxa de Compressão	ssão			
MSE	Rede 1	Rede PCA Adaptativa	otativa	PC	PCA Clássica	ca		JPEG	
	96,875%	93,75%	90,625%	96,875%	93,75%	90,625%	96%	93%	%06
Imagem 1	5.82	4.91	4.27	5.43	4.69	4.21	9.72	7.01	6.10
Imagem 2	6.75	5.49	4.78	6.17	5.17	4.62	9.97	7.58	6.65
Imagem 3	5.20	4.08	3.40	4.64	3.77	3.26	8.91	6.44	5.39
Imagem 4	5.80	4.36	3.58	5.18	4.07	3.44	9.30	6.72	5.64
Imagem 5	6.08	4.50	3.67	5.30	4.16	3.50	9.47	92.9	5.65
Imagem 6	5.19	4.00	3.34	4.64	3.75	3.22	8.90	6.44	5.36
Imagem 7	5.41	3.96	3.33	4.81	3.74	3.21	9.04	6.51	5.39
Imagem 8	5.39	4.06	3.39	4.79	3.81	3.26	8.99	6.48	5.37
Imagem 9	5.32	3.92	3.27	4.66	3.68	3.16	8.93	6.40	5.32
Imagem 10	5.61	4.22	3.53	5.01	3.97	3.41	9.47	6.63	5.57

S S S

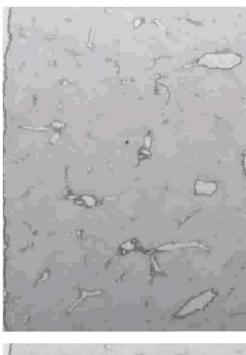
plipl

#### Resultados





Compressão Rede PCA



Compressão JPEG

#### Sumário

- 1 PCA clássica
- Rede PCA adaptativa
- RNA para compressão de imagens
  - Regra de Hebb
- Convergência
- Algoritmo
- Experimentos e resultados
- Conclusões

#### Conclusões

- Três técnicas de compressão foram apresentadas e aplicadas em uma sequência de imagens médicas.
- Analisando os MSE obtidos pelas técnicas:
- Os resultados da Rede PCA foram bem similares aos obtidos pela PCA clássica.
- O desempenho do padrão JPEG foi inferior aos desempenhos obtidos pelas duas outras técnicas.

#### Conclusões

- Analisando a qualidade das imagens recuperadas:
- A compressão JPEG forma padrões que dificultam a análise das imagens.
- As técnicas PCA preservam pequenas estruturas que possibilitam a detecção de doenças.