LABIC

Redes Auto-organizáveis

SCE 5809 REDES NEURAIS

Profa. Roseli Romero

2o. Semestre de 2010

PARTE I

LABIC

PCA

- **Objetivo:** Dadas p variáveis,

$$X_1, X_2, \ldots, X_p$$

- deseja-se achar combinações lineares dessas, para produzir índices que sejam não correlacionados, de tal forma que:
- Índices Z: componentes principais.

LABIC

PCA

i-esima componente principal

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

Com restrição: $a_{i1}^2 + a_{i2}^2 + \cdots + a_{ip}^2 = 1$

$Z_1, Z_2, \ldots, Z_{i-1}, Z_i$ Não correlacionados

LABIC

PCA

- PCA: resume-se em encontrar os autovalores e autovetores da matriz C de covariância dos dados
- C= $\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{bmatrix}$

LABIC

PCA

- λ é auto-valor de A se e somente se existe um vetor não nulo v, tal que A v= λ v.
- Todo vetor que satisfaça a essa condição é um auto-vetor.
- P(λ) = |A - λI| que é um polinômio de grau n em λ é chamado polinômio característico da matriz [a]

LABIC

PCA

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

$$\lambda_1 \geq \lambda_2 \geq .. \geq \lambda_j \geq .. \geq \lambda_p$$

- Os auto-vetores associados:

$$a_1, a_2, \ldots, a_j, \ldots, a_p$$

LABIC

PCA

- Propriedades
- Para $a_i^T a_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$

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

- $a_{i1}, a_{i2}, \ldots, a_{ip}$ são os elementos do *i*-ésimo autovetor correspondente

LABIC

PCA

- a soma dos auto-valores corresponde ao **traco da matriz covariancia C**

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

- var(Z_1) >= var(Z_2) >= … >= var(Z_p)

$$var(Z_i) = \lambda_i$$

LABIC

Exemplo- conj. Iris.dat

		Autovetores (coeficientes)			
Componente	Autovalor	X_1	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	-0.52354

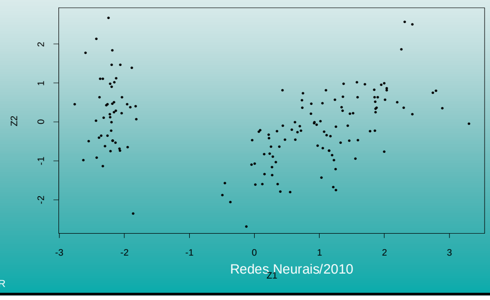
LABIC

PCA

- Conclusões:
- Z1 é responsável por 72.77% do total da variância.
- Z2 é responsável por 23.03% do total da variância.
- Z3 é responsável por 3.68% do total da variância.
- Z4 é responsável por 0.52% do total da variância.

LABIC

PCA



LABIC

Reconstrucao dos dados originais

$$Z = [Z_1, Z_2, \ldots, Z_p]^T$$

$$= [X^T a_1, X^T a_2, \ldots, X^T a_{p-1}]^T$$

$$= A^T X$$

$$X = A\ Z = \sum_{i=1}^{p} Z_i a_i \qquad A^T = A^{-1}$$

LABIC

Reducao da Dimensionalidade

- Sejam $\lambda_1, \lambda_2, \ldots, \lambda_m$
- os m auto-valores da matriz C
- Entao X' ~ X onde:

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

- o erro: e = X - X' onde:

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

LABIC

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

  Verifique isto!!!

  A equação: $e^t X' = 0$: princípio da ortogonalidade
- Existe uma rede neural que implementa PCA, proposta por RUBNER-89.

LABIC

PCA – Conceitos Básicos

- Polinômio Característico

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

  porque $Av = \lambda v \Rightarrow (A - \lambda I)v = 0 \Rightarrow$

  se v não nulo $\Rightarrow \det(A - \lambda I) = 0$

LABIC

Exemplo:

- Determinar os auto-valores e auto-vetores de:

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

$$\det(A - \lambda I) = \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 3 = 0$$

- Sol: $\lambda_1 = 3$ e $\lambda_2 = 1$

LABIC

Exemplo

- Determinar o auto-vetor assoc. a $\lambda_1 = 3$

  $-v_1 + v_2 = 0$

  $v_1 - v_2 = 0$

  Sol: $v_1 = v_2$

  Portanto $v = [1\ 1]^T$

  Analogamente, $v = [1\ \ -1]^T$ para $\lambda_2 = 1$

LABIC

Aplicacao

- **Taxas de Compressão para Blocos 8x8**

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

LABIC

Aplicação

- **Taxas de Compressão para Blocos 16x16**

Número de Componentes Principais	Dimensão do Bloco Compactado	Taxa de Compressão
1	16x1	$1-\left(\frac{1}{16}\right)=93{,}75$
2	16x2	$1-\left(\frac{2}{16}\right)=87{,}5\%$

LABIC

Operações em PCA

- Det. dos auto-valores e auto-vetores:
  calculo do determinante e
  sol. de um sistema de eq. lineares.

LABIC

Exercicio

- Implementar a técnica PCA (use o Matlab ou algum pacote estatístico, ou 1 ling. de progr.) e verifique as 2 comp. Principais obtidas para o conj. iris

LABIC

Reconstrução dos dados originais

$$Z = [\,Z_1, Z_2, \ldots, Z_p\,]^T$$

$$= [\,X^T a_1, X^T a_2, \ldots, X^T a_{p-1}\,]^T$$

$$= A^T X$$

$$X = A\,Z = \sum_{i=1}^{p} Z_i a_i \qquad A^T = A^{-1}$$

LABIC

Redução da Dimensionalidade

- Sejam $\lambda_1, \lambda_2, \ldots, \lambda_m$
- os m auto-valores da matriz C
- Entao X' ~ X onde:

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

- o erro: e = X - X' onde:

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

LABIC

- O vetor de erro e é ortogonal ao vetor X', que aproxima X.
  Verifique isto!!!
  A equação: $e^T X' = 0$: princípio da ortogonalidade
- Existe uma rede neural que implementa PCA, proposta por RUBN-89 (vamos ver!!!)

LABIC

•Algoritmo da PCA

1. Padronize os dados correspondentes às variáveis para que estes tenham média igual a 0 e variância igual a 1.
2. Calcule a matriz de covariância *C*.
3. Encontre os autovalores da matriz *C* e seus correspondentes auto-vetores. Os coeficientes da *i*-ésima componente principal são dados pelo auto-vetor associado ao i-ésimo auto-valor

LABIC

4. Descarte as componentes que acumulem uma pequena proporção da variação dos dados. Por exemplo, se os dados originais tiverem 20 variáveis e as três primeiras componentes principais forem responsáveis por 90% do total da variância, as outras 17 componentes principais podem ser ignoradas.

LABIC

Aplicacao

- **Taxas de Compressão para Blocos 8x8**

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

LABIC

Aplicação

- **Taxas de Compressão para Blocos 16x16**

Número de Componentes Principais	Dimensão do Bloco Compactado	Taxa de Compressão
1	16x1	$1-\left(\frac{1}{16}\right)=93{,}75$
2	16x2	$1-\left(\frac{2}{16}\right)=87{,}5\%$

LABIC

Algoritmo PCA - Matlab

```
%%%%%%%%%%%%%%%%%%
%  Componentes Principais      %
%     PCA  - Roseli Romero     %
%%%%%%%%%%%%%%%%%%
% carregamento doa arquivo iris.dat
load iris.txt
% verificando a dimensao do conjunto
[n,p] = size(iris);
% tirando a primeira coluna da matriz de dados
iris(:,1)=[ ];
p= p-1;
X = iris;   % matriz X contem os dados
% centrando os dados na media 0
S = std(X);  % S armazena os desvios-padrao de cada coluna de X
M = mean(X); % M armazena as medias de cada coluna de X
X = X - ones(n,1) * M;
% transformando os dados para ter variancia 1
X = X./(ones(n,1)*S);
% calculando a matriz de covariancia C
C = (X'*X)/n;
% auto-valores (A) e auto-vetores (V) da matriz de covariancia
[V,A] = eig(C);
A = diag(A);
```


LABIC

```
% ordenando auto-valores e auto-vetores por ordem crescente de auto-valores
V = V';   %coloca auto-vetores nas linhas de V
A= [A,V]; % concatena A e V (cada linha de A contem um auto-valor e auto-vetor)
A=sortrows(A,1);  %ordena os auto-valores em ordem crescente
A=A(:,1) %separa auto-valores em A e imprimi na tela
V       %imprimi V na tela
%calculando as componentes principais
Z=[];
for i=1:p
  Vt = V(p-i+1,:); %pega os auto-vetores em ordem decrescente dos auto-valores
  Z(:,i) = X * Vt';  % otem-se a i-esima componente principal
end
% salvando as componentes principais em um arquivo
save('componentes_iris.txt')
Z1=Z(:, 1:2)
L= Z1(1:50,:);
L1=Z1(51:100,:);
L2=Z1(101:150,:);
% fazendo o grafico das duas primeiras componentes
%plot(Z1,Z2,'-')
%fazendo o grafico das tres classes separadas
plot(L(:,1),L(:,2),'.',L1(:,1),L1(:,2),'+',L2(:,1),L2(:,2),'*')
```

LABIC

RNA para Compressão de Imagens

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

LABIC

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

LABIC

RNA PCA Adaptativa

- Postulado de Hebb[Hebb-1949]: "Quando um axônio da célula A está suficientemente próximo para excitar uma célula B e repetidamente tenta excitá-la, algum processo crescente ou mudanças metabólicas ocorrem em ambas as células".
- Transformando em regras [Sten-73]:

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

LABIC

- Regra de Hebb:

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

Onde η é uma constante positiva – veloc.de aprendizado

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

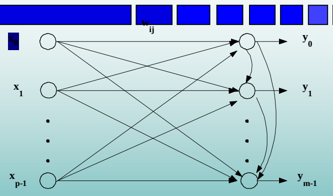
x_j é j-ésimo elemento do vetor de entrada no tempo n

- Regra Anti-Hebbiana [Foldiak, 1989]

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

LABIC

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: $\Delta w_{ij}(n) = \eta x_i(n) y_j(n)$

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

LABIC

Teorama de Convergência[Sang-89]

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

LABIC

Teorema

Portanto, no limite, pode-se escrever:

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

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

LABIC

- Pode-se acelerar a convergência da rede introduzindo um termo *momentum* β e deixando que os parâmetros de aprendizagem e o *momentum* diminuam com o tempo. Em [MAO-95], utilizou-se:

$$\Delta w_{ij}(n+1) = \eta(n) x_i y_j + \beta(n) \Delta w_{ij}(n) \qquad (1)$$

LABIC

$$\Delta u_{lj}(n+1) = -\mu(n) y_l y_j + \beta(n) \Delta u_{lj}(n) \qquad (2)$$

onde $\eta(n+1) = \max(\alpha\eta(n), 0.0001)$, $\mu(n+1) = \max(\alpha\mu(n), 0.0002)$, $\beta(n+1) = \max(\alpha\beta(n), 0.0001)$ e α é o fator de diminuição.

LABIC

Algoritmo PCA Adptativa-[Rubn-89]

Inicio

1. Inicialize todos os pesos de conexões com pequenos valores aleatórios e escolha os valores para os parâmetros de aprendizagem.Normalize-os em [0,1]. Se normalizar em [-1,1] pode mudar os sinais dos auto-vetores.

2. Repita

2.2 Selecione aleatoriamente um padrão *p*-dimensional e apresente-o à rede.

2.3. Ajuste os pesos das conexões entre a camada de entrada e a camada de saída de acordo com a Eq. (1)

2.4. Normalize os vetores-peso (em colunas).

2.5. Atualize os pesos laterais de acordo com a equacao (2)(não precisa normalizar)

2.6. Modifique os parâmetros β, η, e μ.

Até que {todos os pesos laterais sejam suficientemente pequenos (a soma de seus valores absolutos seja menor que algum *threshold* ε)} ou {um número de iterações máximo seja atingido}.

Fim.

LABIC

Experimentos Realizados

- 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

LABIC

Taxas de Compressão

$$Taxa\ de\ Compressão = 1 - \frac{comprimento\ da\ cadeia\ de\ dados\ comprimidos}{comprimento\ da\ cadeia\ de\ dados\ originais}$$

Taxas de Compressão para Blocos 32x32

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

LABIC

Compressão de Imagens através do Padrão JPEG

- Utilização do aplicativo xv (xview) para linux:
  - possui módulo para compressão JPEG
  - possibilita que o usuário selecione taxa de compressão

LABIC

Discussão de Resultados (1)

MSE	Taxa de Compressão								
	Rede PCA Adaptativa			PCA Clássica			JPEG		
	96,875%	93,75%	90,625%	96,875%	93,75%	90,625%	96%	93%	90%
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	6.76	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

LABIC

Discussão de Resultados (2)

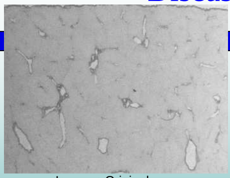
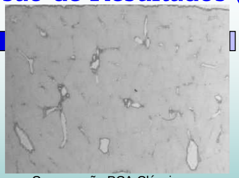
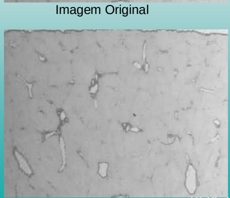


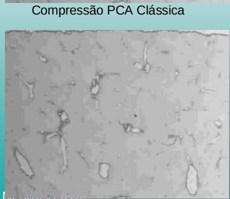
Imagem Original



Compressão PCA Clássica



Compressão Rede PCA



Compressão JPEG

LABIC

Conclusões

- 3 técnicas de compressão foram apresentadas e aplicadas em uma sequência de imagens médicas
- Analisando os MSE cometidos 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

LABIC

Conclusões (cont.)

- 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