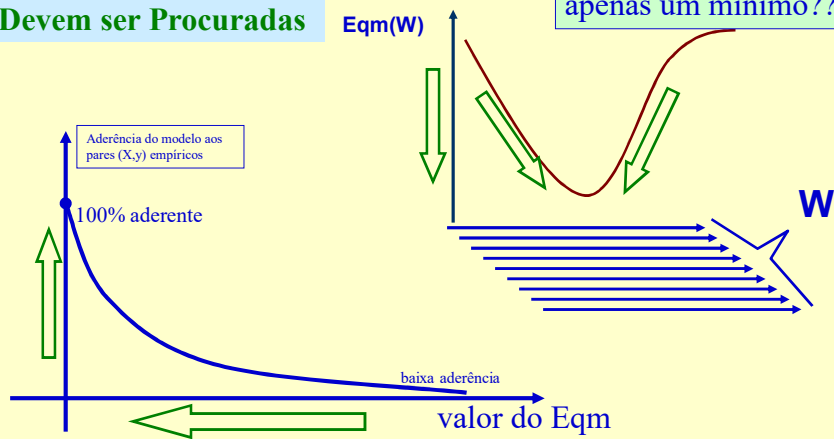


O que devemos mirar quando exploramos o espaço de pesos W buscando que a RNA seja um bom modelo?

Devemos buscar Maximização da aderência = Mínimo Eqm possível

As Setas Verdes Indicam Situações que Devem ser Procuradas



135

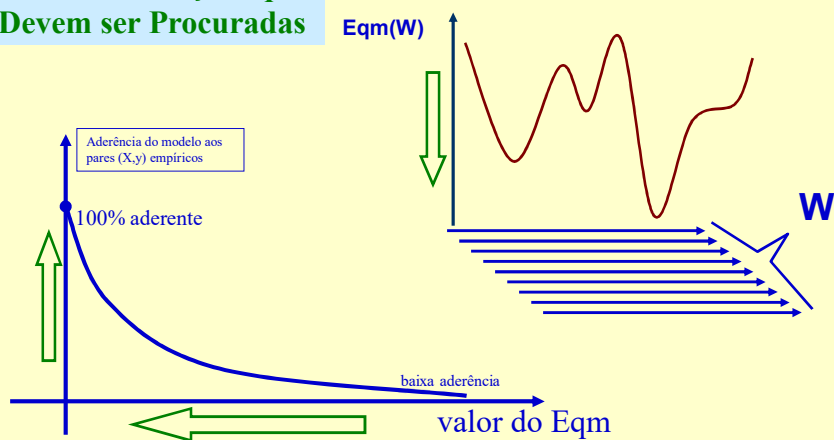
© Prof. Emilio Del Moral – EPUSP

135

O que devemos mirar quando exploramos o espaço de pesos W buscando que a RNA seja um bom modelo?

Devemos buscar Maximização da aderência = Mínimo Eqm possível

As Setas Verdes Indicam Situações que Devem ser Procuradas

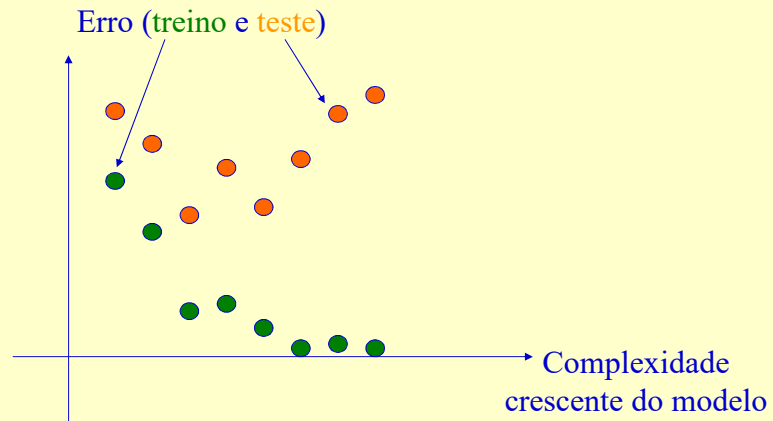


136

© Prof. Emilio Del Moral – EPUSP

136

Atenção para componentes randômicas que impactam muito quando se faz um único ensaio de medida de erro, para cada tamanho de rede específico (um ensaio apenas, para cada grau de complexidade) ...



137

© Prof. Emilio Del Moral – EPUSP

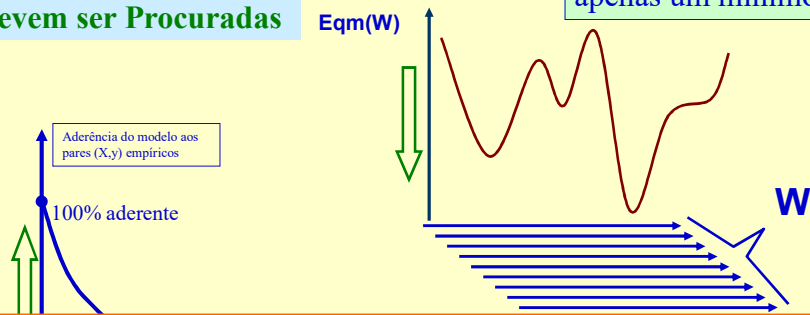
137

O que devemos mirar quando exploramos o espaço de pesos W buscando que a RNA seja um bom modelo?

Devemos buscar Maximização da aderência = Mínimo E_{qm} possível

As Setas Verdes Indicam Situações que Devem ser Procuradas

Será que temos apenas um mínimo??

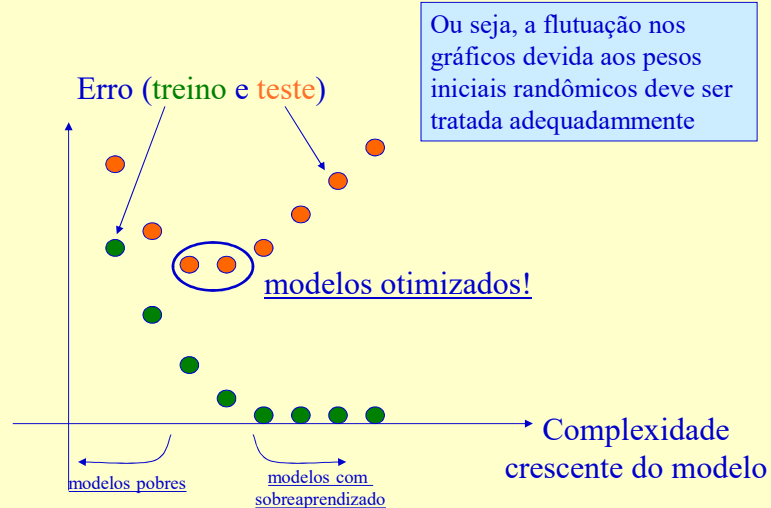


... Para não sermos reféns de mínimos locais com alto E_{qm} , podemos aplicar o gradiente descendente repetidamente na mesma RNA, com novos pesos iniciais randômicos em cada rodada, mantendo para o modelo final apenas os valores de pesos associados ao ensaio com o melhor dos resultados finais no E_{qm} !

© Prof. Emilio Del Moral – EPUSP

138

Com repetidos ensaios em cada grau de complexidade os mínimos locais são evitados e detectamos adequadamente o sobreaprendizado



139

© Prof. Emilio Del Moral – EPUSP

139

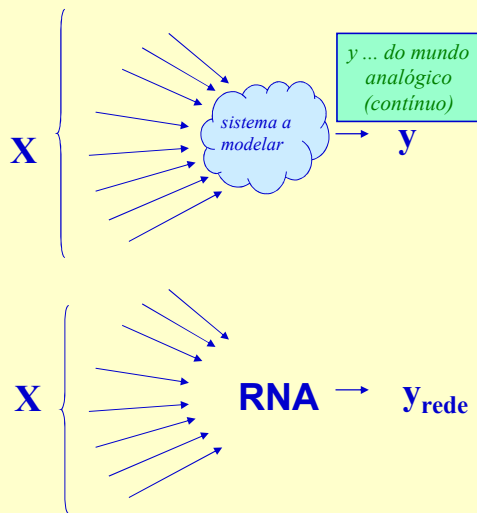
Identificando os ingredientes para o risco de sobreaprendizado nos contextos de regressão multivariada e de reconhecimento de padrões multivariado

144

© Prof. Emilio Del Moral – EPUSP

144

**Modelagem de um sistema por função de mapeamento $X \rightarrow y$
(a RNA como regressor analógico não linear multivariável)**



Assumimos que a variável y do sistema a modelar é uma função (normalmente desconhecida e possivelmente não linear) de diversas outras variáveis desse mesmo sistema

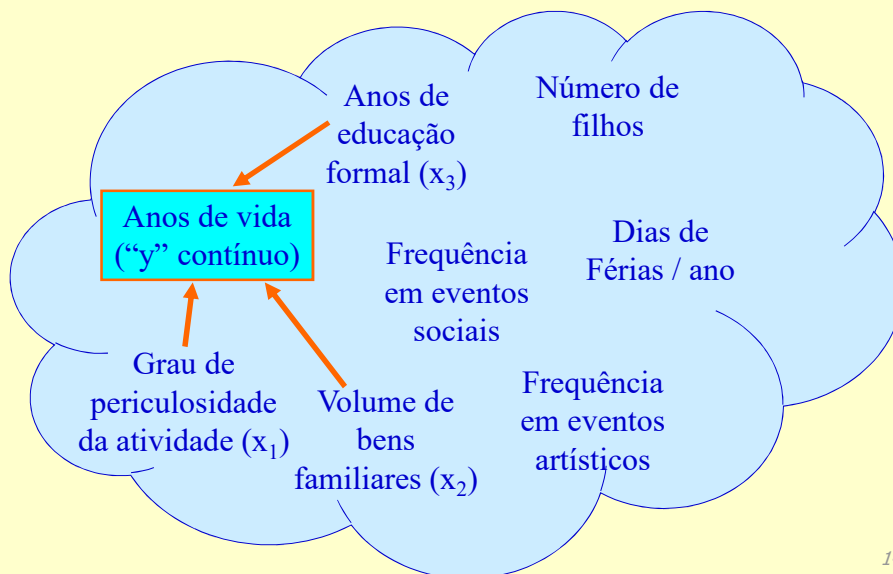
A RNA, para ser um bom modelo do sistema, deve reproduzir essa relação entre X e y , tão bem quanto possível

145

© Prof. Emilio Del Moral – EPUSP

145

Um hipotético universo de variáveis interdependentes, passível de modelagem/ens

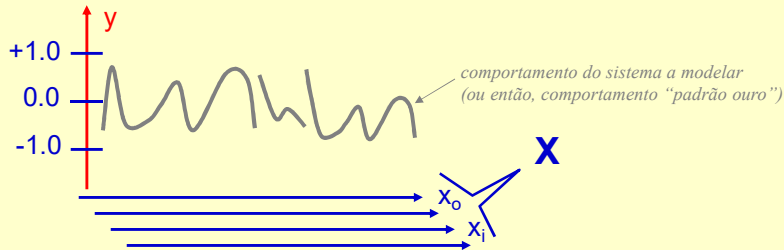


146

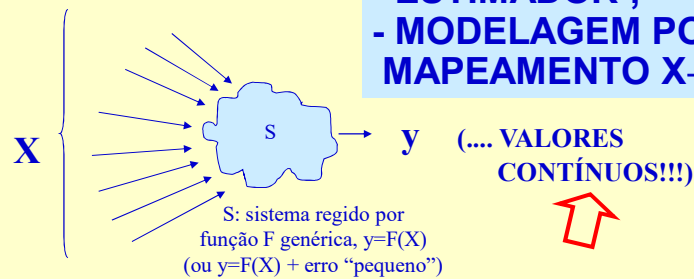
© Prof. Emilio Del Moral – EPUSP

146

A função $y(X)$ “a descobrir”, num caso geral de função analógica $y(X)$



- ESTIMADOR ;
- MODELAGEM POR
MAPEAMENTO $X \rightarrow y$

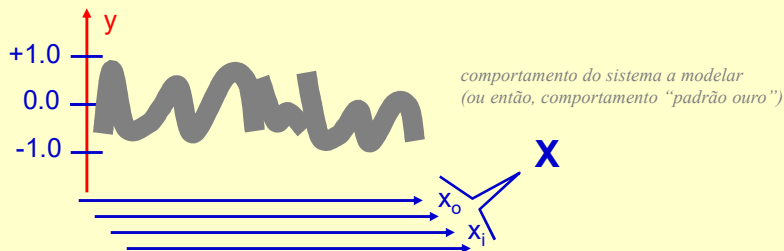


147

© Prof. Emilio Del Moral – EPUSP

147

Cenário mais real: a “função” $y(X)$ do sistema modelado é “difusa”: $y = F_{\text{médio}}(X) + \text{flutuação}$



.... em problemas concretos / reais, há sempre alguma ambiguidade no mapeamento que leva valores de X a valores de y . Para decepção de Cybenko, não temos uma função $y=F(X)$ no sentido matemático exato, pois para uma dada ênupla de valores X fixados, temos tipicamente uma faixa de valores que podem ser observados para a variável y : $y = F_{\text{médio}}(X) + \text{flutuação}$.

Neste cenário, buscamos que o modelo capture o comportamento médio das relações observadas entre X e y :

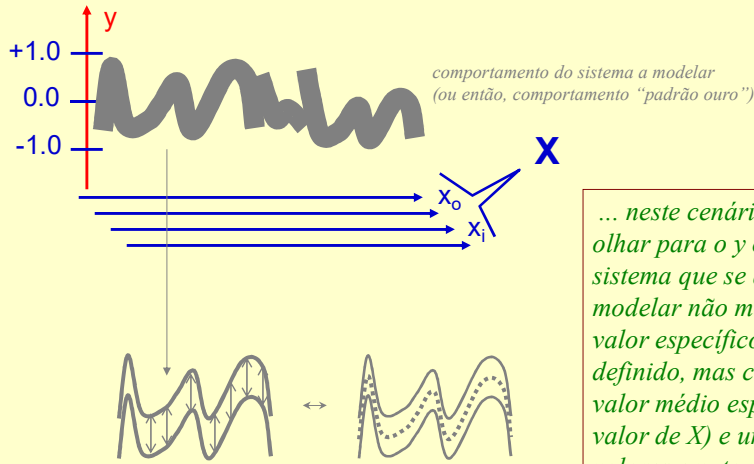
$$\dots y_{\text{rede}} \sim y_{\text{médio}} \text{ esperado para um dado } X$$

148

© Prof. Emilio Del Moral – EPUSP

148

Cenário mais real: a “função” $y(X)$ do sistema modelado é “difusa”: $y = F_{\text{médio}}(X) + \text{flutuação} \dots$



... neste cenário, podemos olhar para o y observado no sistema que se deseja modelar não mais como um valor específico bem definido, mas como um valor médio esperado (dado valor de X) e uma faixa de valores em torno desse valor médio esperado.

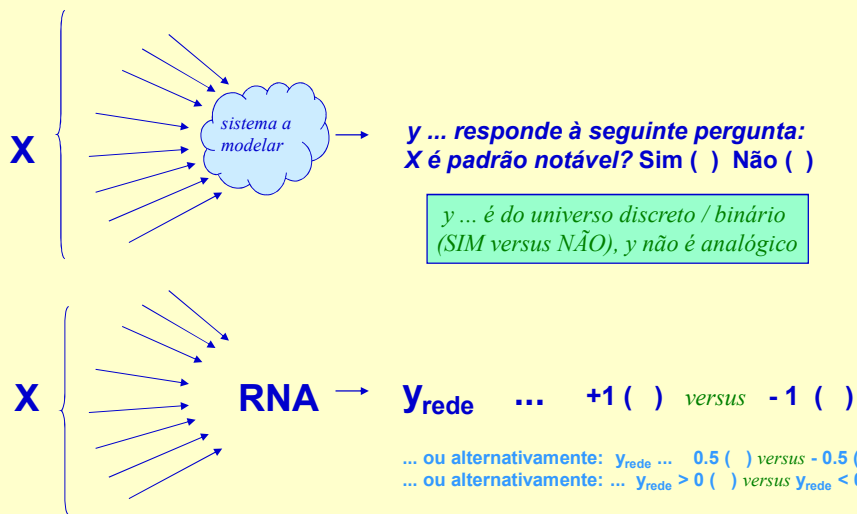
149

© Prof. Emilio Del Moral – EPUSP

149

RNAs como reconhecedor / detetor de padrões

...

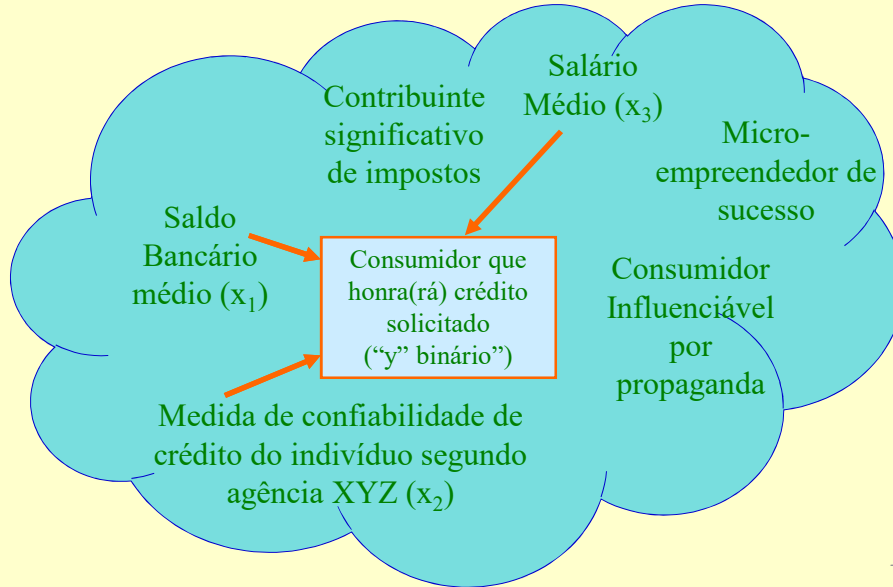


150

© Prof. Emilio Del Moral – EPUSP

150

Um hipotético universo de variáveis interdependentes, passível de modelagem/ens

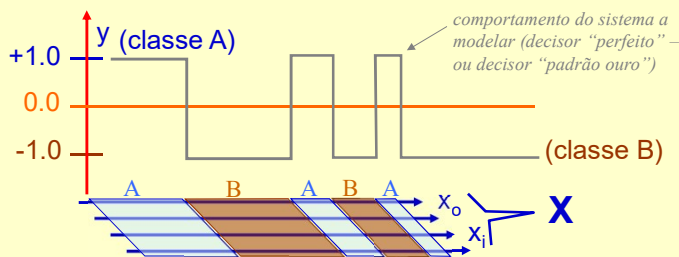


151

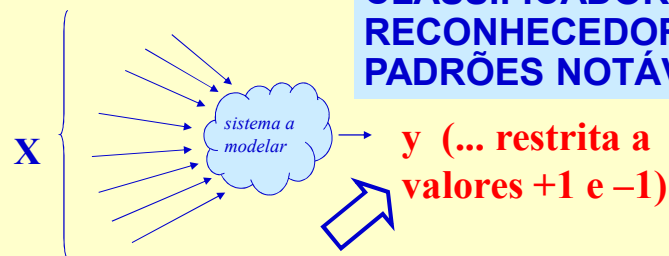
© Prof. Emilio Del Moral – EPUSP

151

Caso de classificação binária / reconhecimento de padrões, será do tipo ...



**CLASSIFICADOR;
RECONHECEDOR DE
PADRÕES NOTÁVEIS**



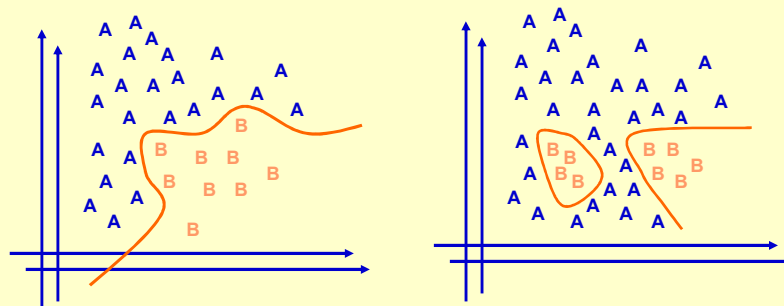
152

© Prof. Emilio Del Moral – EPUSP

152

Capacidade de reconhecimento de padrões em casos complexos NÃO LINEARES

Com as RNAs, a hipersuperfície de separação entre classes vai muito além dos hiperplanos

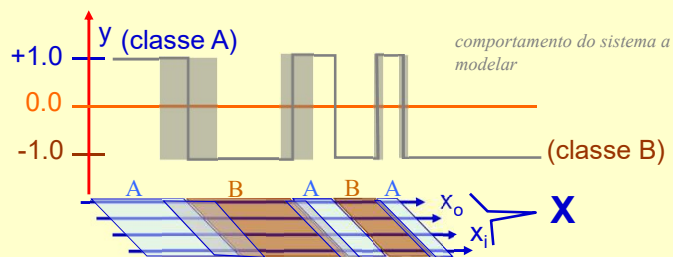


153

© Prof. Emilio Del Moral – EPUSP

153

Cenário mais real: a separação entre regiões do espaço de X não é perfeitamente definida



.... em problemas concretos / reais, há sempre alguma ambiguidade no mapeamento que leva valores de X aos valores discretos de y . Não temos uma função $y=F(X)$ no sentido matemático exato, pois para uma dada ênupla de valores X fixado temos em alguns casos de fronteira a possibilidade de observar no y empírico tanto a classe A quanto a classe B : $y=A$ ou B , com maior ou menor probabilidade para cada classe de acordo com o X . Neste desejamos que o modelo capture o comportamento médio das relações observadas entre X e y :

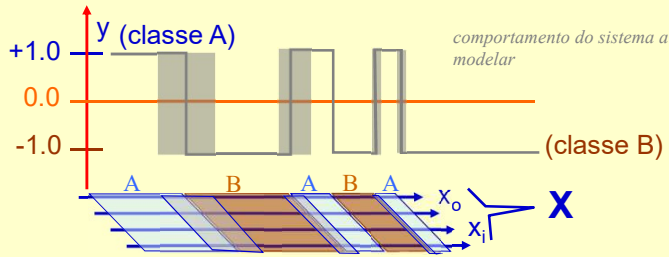
... $y_{rede} \sim$ classe 'mais esperada' para um dado X

154

© Prof. Emilio Del Moral – EPUSP

154

Cenário mais real: a separação entre regiões do espaço de X não é perfeitamente definida



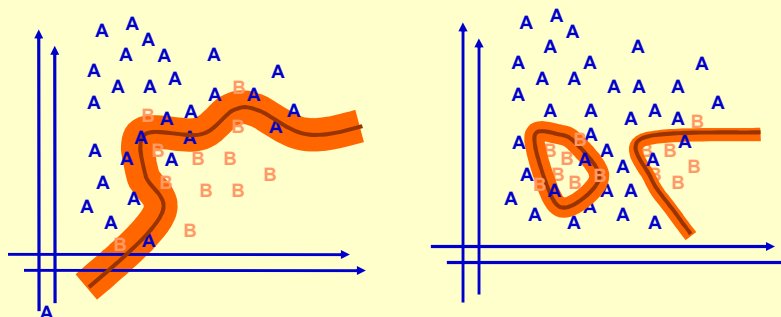
... podemos olhar para o y (classe A ou B) observado no sistema que se deseja modelar não mais como uma classe sempre bem definida e com fronteiras de separação entre A e B bem definidas no espaço de valores de X , mas como sendo delineadas na modelagem através de fronteiras com eventuais faixas de tolerância e com sobreposição parcial das classes no espaço de X

155

© Prof. Emilio Del Moral – EPUSP

155

Situações de classes com sobreposição parcial no espaço de atributos X ; situações de fronteiras de separação difusas ...



157

© Prof. Emilio Del Moral – EPUSP

157

O Ciclo completo da modelagem:

0) Formalização do problema, mapeamento quantitativo em um modelo neural inicial e ... 0b) coleta de pares empíricos (X,y)

1) Fase de **TREINO** da RNA (MLP): com conhecimento dos X e dos y, que são ambos usados na calibração do modelo

2) Fase de **TESTE** / Caracterização da qualidade da RNA para generalizar: temos novos pares X e y, com y guardado "na gaveta", usado apenas para avaliação, não para re-calibração. É como um ensaio de uso final do modelo, com possibilidade de medir a sua qualidade com o y que foi guardado na gaveta.

[Fase de refinamentos sucessivos da RNA e/ou dos dados e/ou do modelo, em ciclos diversos, recomeçando desde o passo 0 ou do passo 1]

3) Fase de **USO FINAL** da RNA, com y efetivamente não conhecido, e estimado com conhecimento dos X + uso do modelo calibrado.

.... Diferenças e semelhanças entre 1, 2 e 3

163

© Prof. Emilio Del Moral – EPUSP

163

164

Revisitando os Conjuntos de Dados Empíricos ...

Treino + Teste ...

versus

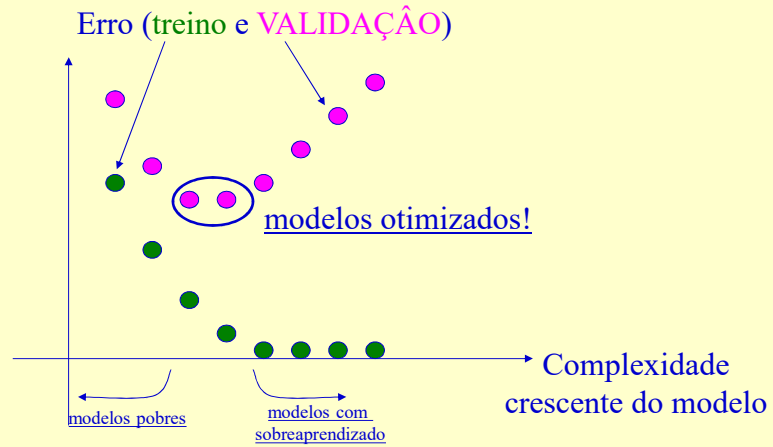
Treino + **Validação** + Teste

© Prof. Emilio Del Moral Hernandez

164

164

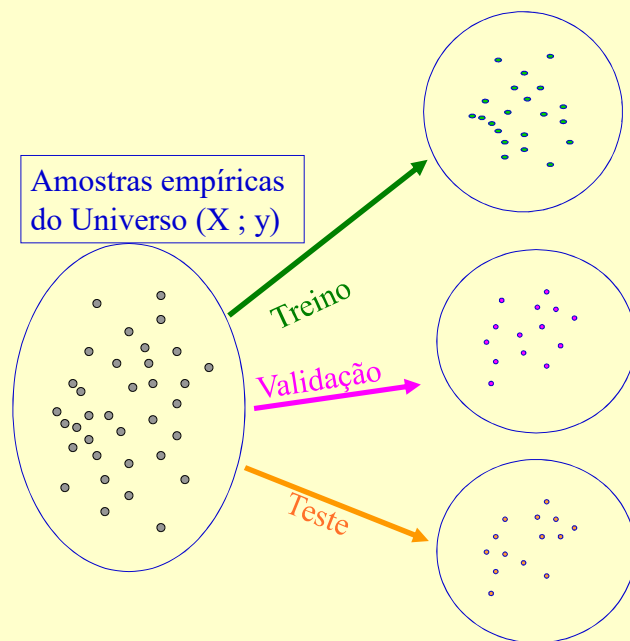
Sobreaprendizado em “sumário executivo”



165

© Prof. Emilio Del Moral – EPUSP

165

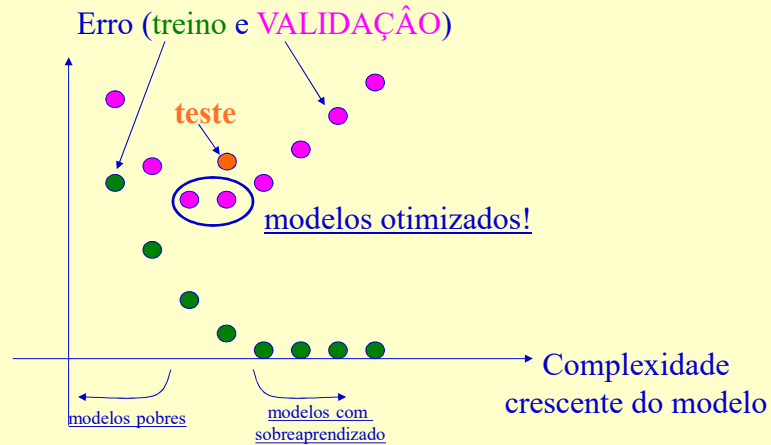


166

© Prof. Emilio Del Moral – EPUSP

166

Sobreaprendizado em “sumário executivo”



167

© Prof. Emilio Del Moral – EPUSP

167

Alguns dos diversos usos do conceito de conjunto de validação, adicional ao conjunto de treino e de teste:

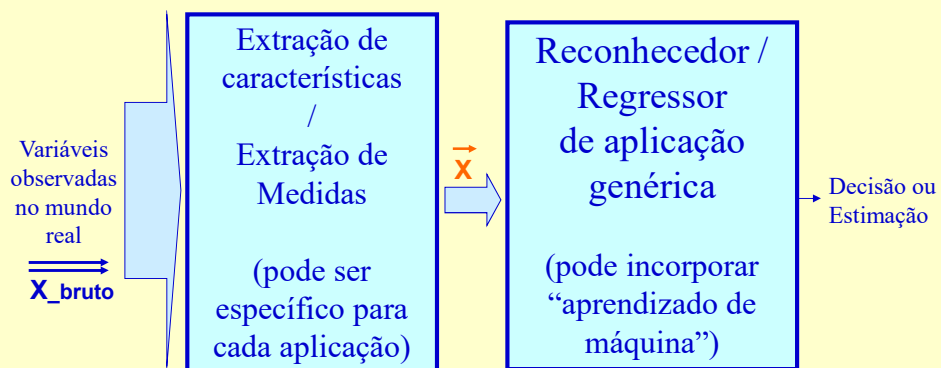
- *Seleção de complexidade do modelo neural para limitação de sobreaprendizado*
- *Ativação do early stop no aprendizado (Matlab): critério de parada adicional no processo de refinamento de pesos sinápticos*
- *Balizador no processo de seleção de estratégias de pré-processamento / extração de medidas X alternativos*
- *... etc*

168

© Prof. Emilio Del Moral – EPUSP

168

Elaborando uma Solução em dois estágios



169

© Prof. Emilio Del Moral – EPUSP

169

172

© Prof. Emilio Del Moral – EPUSP

172

Vimos em tópico anterior uma estratégia de refinamento de um protótipo de regressão / reconhecimento através da sintonia do número de neurônios / complexidade da RNA, para bom compromisso entre precisão do modelo (que pede número alto de neurônios) e controle de sobreaprendizado (que pede número baixo de neurônios).

Falemos agora de um outro aspecto que pode ser considerado no ciclo genérico de refinamentos do regressor / do reconhecedor:

O pré-processamento, a reescrita e a redução de dimensionalidade do vetor de entradas X do modelo

174

© Prof. Emilio Del Moral – EPUSP

174

O Ciclo completo da modelagem:

0) **Formalização do problema, mapeamento quantitativo em um modelo neural inicial e ... 0b) coleta de pares empíricos (X,y)**

1) **Fase de TREINO da RNA (MLP): com conhecimento dos X e dos y, que são ambos usados na calibração do modelo**

2) **Fase de TESTE / Caracterização da qualidade da RNA para generalizar: temos novos pares X e y, com y guardado “na gaveta”, usado apenas para avaliação, não para re-calibração. É como um ensaio de uso final do modelo, com possibilidade de medir a sua qualidade com o y que foi guardado na gaveta.**

[Fase de refinamentos sucessivos da RNA e/ou dos dados e/ou do modelo, em ciclos diversos, recomeçando desde o passo 0 ou do passo 1]

3) **Fase de USO FINAL da RNA, com y efetivamente não conhecido, e estimado com conhecimento dos X + uso do modelo calibrado.**

.... **Diferenças e semelhanças entre 1, 2 e 3**

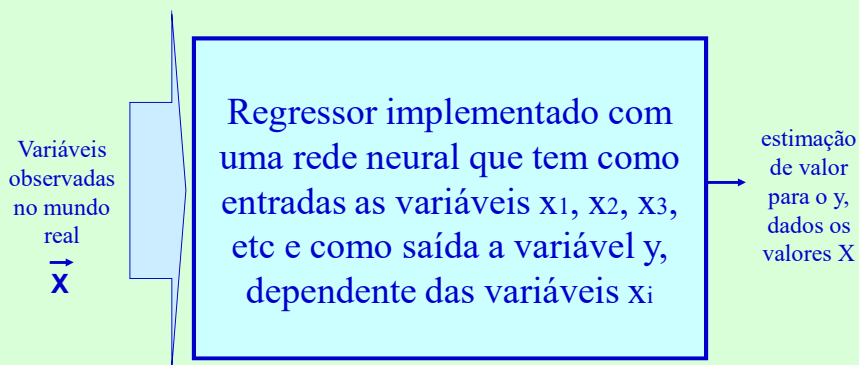
175

© Prof. Emilio Del Moral – EPUSP

175

Regressor multivariado

177



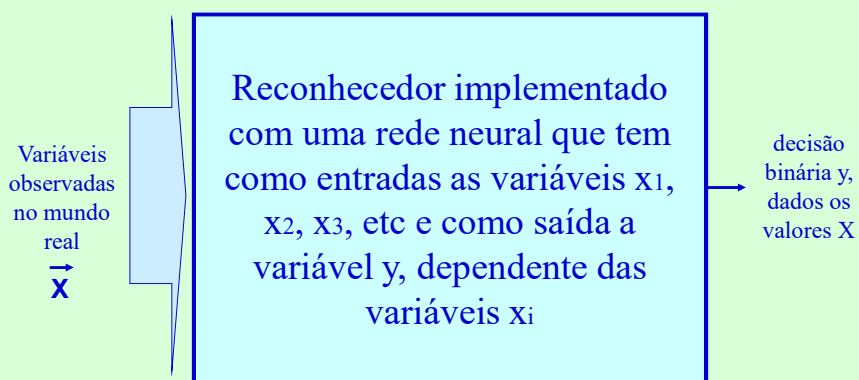
© Prof. Emilio Del Moral Hernandez

177

177

Detector / reconhecedor multivariado

178



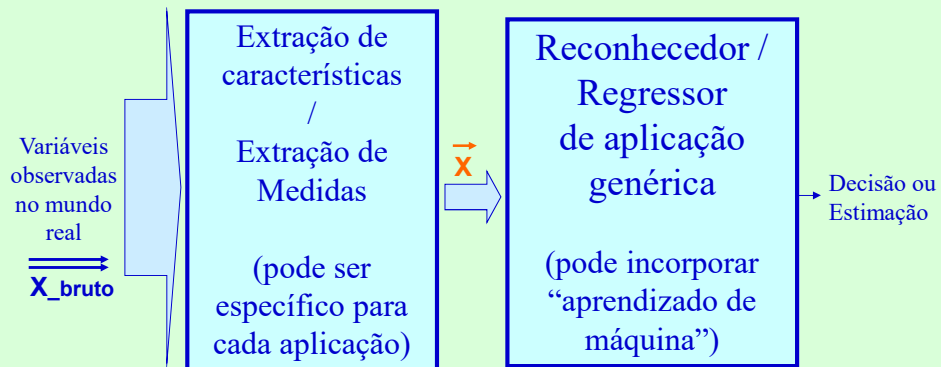
© Prof. Emilio Del Moral Hernandez

178

178

Elaborando uma Solução em dois estágios

181



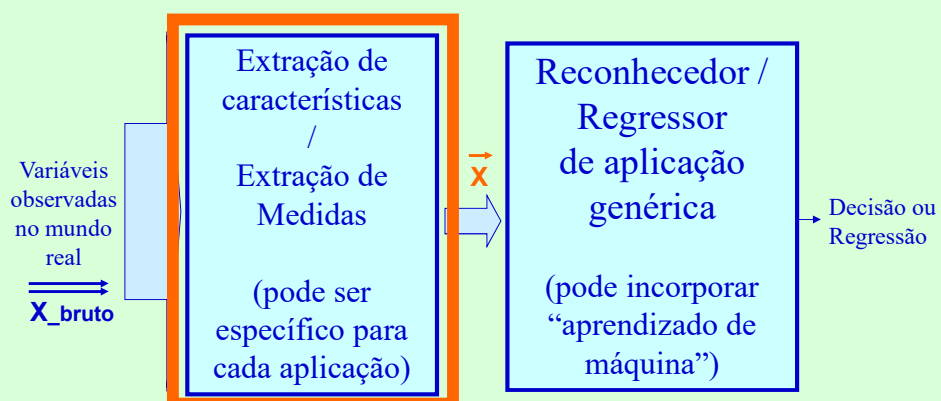
© Prof. Emilio Del Moral Hernandez

181

181

... O 1o estágio gera um Vetor de Medidas, \vec{x} (o segundo estágio operará sobre tal vetor)

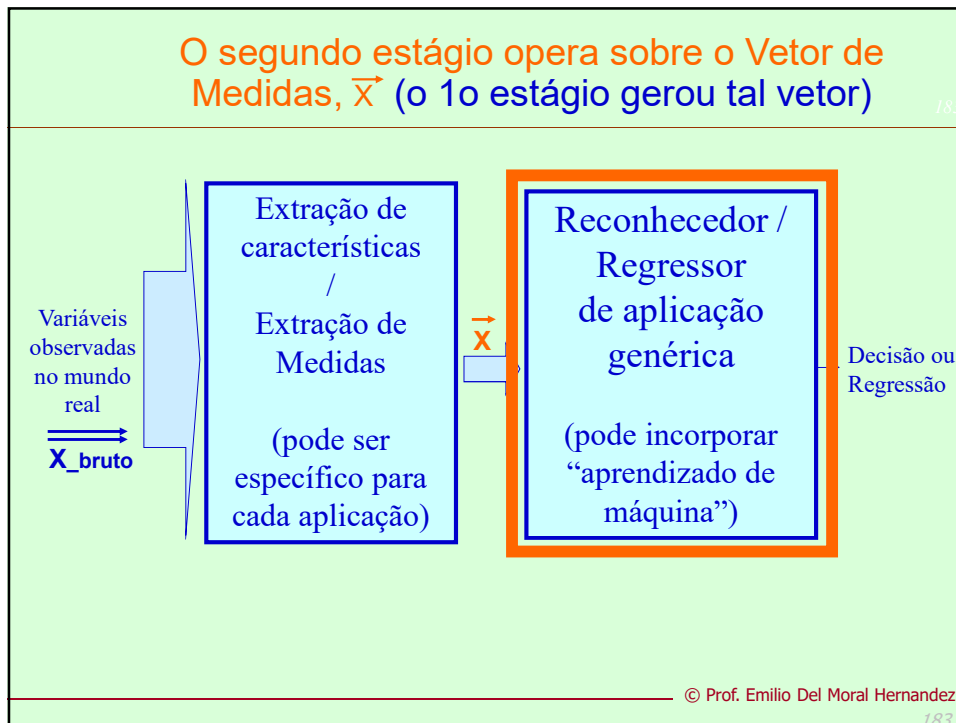
182



© Prof. Emilio Del Moral Hernandez

182

182



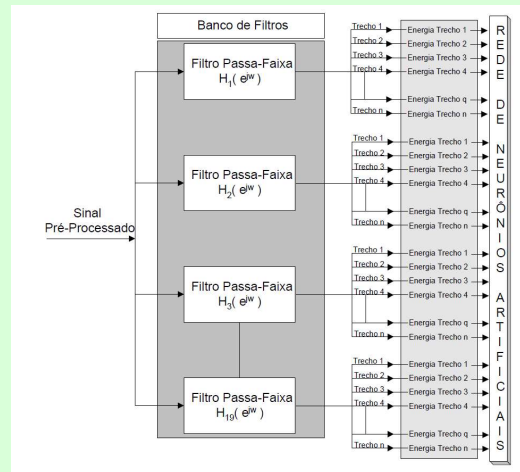
183

- Alguns exemplos de grandezas componentes dos vetores de medidas X:**
- Grandezas macroscópicas como energia do sinal, amplitude, frequência média ...
 - Componentes de diversas harmônicas (análise em frequência)
 - Componentes de análise tempo-frequência
 - Intensidades luminosas ou intensidades em canais de cor (RGB por ex.)
 - Histogramas de intensidades
 - Principal Components (componentes principais – PCA)
 - Medidas sobre séries temporais (médias móveis, por exemplo; medidas de dispersão / instabilidade localizadas)
 - Medidas específicas à aplicação, experimentadas em problemas similares ao seu, relatadas na literatura técnica como sendo de sucesso
- 184
- © Prof. Emilio Del Moral Hernandez
184

184

TCC em Reconhecimento de comandos de voz (Daniel e Gabriel) – sobe, desce, esquerda, direita

185



© Prof. Emilio Del Moral Hernandez

185

185

Medida de qualidade do reconhecedor de comandos de voz ensaiado ...

186

amostra \ resultado	SOBE	DESCE	ESQUERDA	DIREITA
SOBE	85%	0%	15%	0%
DESCE	0%	100%	0%	0%
ESQUERDA	0%	0%	80%	20%
DIREITA	0%	0%	10%	90%

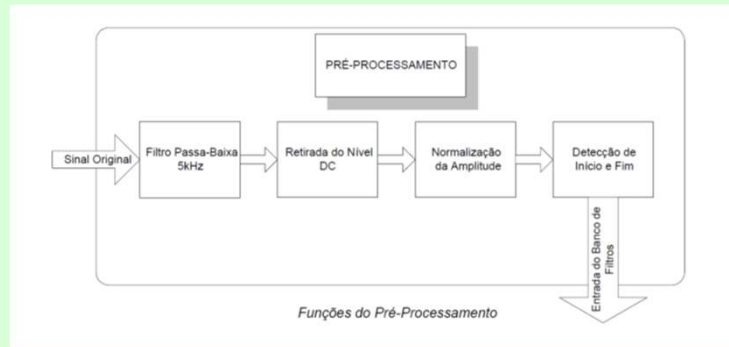
© Prof. Emilio Del Moral Hernandez

186

186

Exemplo de pré-processamentos num reconhecedor de comandos de voz - TCC de Daniel e Gabriel

187



© Prof. Emilio Del Moral Hernandez

187

187

Exemplo de modelagem / reconhecimento automático de padrões ... (tese de Liselene / Prof Miguel)

188

O procedimento de detecção do Vazamento

- * Haste de escuta
- * Correlacionador de ruído
- * Geofone

Três fotografias em sequência mostram um técnico em um ambiente externo realizando procedimentos de detecção de vazamentos. Na primeira foto, o técnico utiliza uma haste de escuta para investigar o chão. Na segunda, utiliza um geofone conectado a um equipamento eletrônico. Na terceira, utiliza um correlacionador de ruído para identificar fontes de vazamento.

© Prof. Emilio Del Moral Hernandez

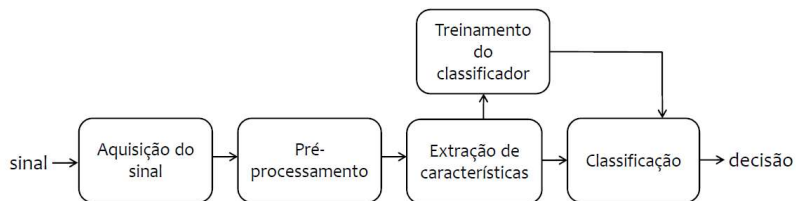
188

188

Exemplo de modelagem / reconhecimento automático de padrões ... (tese de Liselene / Prof Miguel)

189

Sistema de Detecção de Vazamento



© Prof. Emilio Del Moral Hernandez

189

189

Exemplos de projetos concebidos e realizados por alunos da disciplina PSI-2672

190

The screenshot shows a web browser window displaying a page titled "PSI 2672 - Práticas em Reconhecimento de Padrões, Modelagem e Neurocomputação". The page lists several final projects:

- 1. Classificação em gêneros musicais ([gr_I_apresentacao_Musicas.pdf](#))
- 2. Implementação de um sistema para correção da medida de um sensor de pressão diferencial ([gr_II_Apresentacao_tna_sensor_pressao_03_01.pdf](#))
- 3. Análise da qualidade dos vinhos a partir de testes físicos-químicos ([gr_III_apresentacao_PSI2672 - Análise de vinhos.pdf](#))
- 4. Reconhecimento de placas de trânsito ([gr_IV_apresentacao_Reconhecimento_de_placas.pdf](#))
- 5. Reconhecedor de alcoolismo ([gr_V_apresentacao_Reconhecedor Alcoolismo.pdf](#))
- 6. Língua eletrônica ([gr_VI_apresentacao_lingua_eletronica.pdf](#))
- 7. Reconhecimento de dígitos ([gr_VII_apresentacao_Reconhecedor de digitos.pdf](#))

At the bottom of the page, there is a link for "Informações sobre o programa do curso: [Programa PSI2672_2011.pdf](#)".

© Prof. Emilio Del Moral Hernandez

190

190

Classificação automática de generos musicais

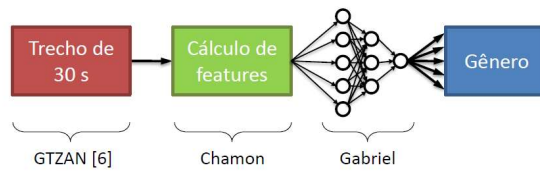
193

DESCRIÇÃO DO PROBLEMA

Categorizar músicas em gêneros automaticamente baseado em trechos de suas gravações e exemplos rotulados *a priori*

- 1) Blues
- 2) Classical
- 3) Country
- 4) Disco
- 5) Hip-Hop
- 6) Jazz
- 7) Metal
- 8) Pop
- 9) Reggae
- 10) Rock

A SOLUÇÃO



projeto de alunos em PSI-2672

© Prof. Emilio Del Moral Hernandez

193

193

Seguidor do alvo dos olhos na tela do computador

194

Eye-tracker – Definição do problema

"Utilizar técnicas neurais para determinar a posição (x,y) para a qual os olhos do usuário estão apontados na tela"



Coleta de dados



Método semi-automático de coleta de dados:

- Divisão da tela do computador em 25 quadrantes
- Para cada quadrante, tirar 20 fotos do olho (somente do olho, não do rosto todo!) apontando para a posição (com ajuda de software para tirar fotos e salvá-las)

Total de elementos da amostra: 500 fotos

Pré-processamento



projeto de alunos em PSI-2672

194

194

Curso: PSI3471 - Funda x + v

https://edisciplinas.usp.br/course/view.php?id=66581

TAREFA #4 para entrega em escaninho e-disciplinas até dia 03 de abril (escaninho aberto no dia 28/março):

- Apoio para atividade em sala de 27-março-2019 com entrega para nota
- Enunciado postado em forma definitiva em 28-março fim do dia
- Entregue aqui a tarefa #4 - Derivação analítica + Cálculo de incrementos de 3 pesos sinápticos e avaliação de melhoria do erro quadrático + Tempo da tarefa

----- // -----

- Registro aproximado de temas aula a aula, ate a aula #11 em 01-04-2019

Exemplos de sistemas inteligentes usando técnicas de regressão multivariada e/ou reconhecimento multivariado (materiais usados como exemplos nas aulas do Prof Emilio)

- Lista / Inventário com Projetos PSI 2672 (disciplina prática de 5o ano) de 2011 ate 2016
- Repositório (36Mb) de apresentações finais de PSI2672 (disciplina de 5o ano) nos anos 2011 a 2016 – perto de 50 projetos de alunos de 5o ano / Prof Emilio Del Moral Hernandez
- RESUMO DAS APRESENTAÇÕES FINAIS DOS PROJETOS DE PSI3571 EM 2018 (Disciplina de 5o ano)
- RESUMO DAS APRESENTAÇÕES FINAIS DOS PROJETOS DE PSI2672 EM 2017 (Disciplina de 5o ano)
- Vitrine com 10 apresentações finais de PSI5886-2018 (Princípios de Neurocomputação - Pós Grad); apresentações finais dos 10 grupos de alunos que cursaram a disciplina em 2018

© Prof. Emilio Del Moral Hernandez

195

195

ONE

USP

PCA – Principal Components Analysis

Teoria e Aplicações

Prof. Dr. Edson C. Kitani
Prof. Dr. Emilio Del Moral Hernandez

199









GRUPO DE INTELIGÊNCIA COMPUTACIONAL, MODELAGEM E NEUROCOMPUTAÇÃO ELETRÔNICA – ICONE-EPUSP

LABORATÓRIO DE SISTEMAS INTEGRADOS –LSI
DEPARTAMENTO DE ENGENHARIA DE SISTEMAS ELETRÔNICOS – PSI
ESCOLA POLITÉCNICA DA USP

200






Prof. Dr. Edson C. Kitani

<http://buscatextual.cnpq.br/buscatextual/visualizacv.do?id=K4269320J9>

Formação:
 Técnico em Eletrônica (ETEP – Santo André)
 Tecnólogo em Automação Industrial (Centro Universitário de Santo André)
 Especialista em Mecânica Fina (Universidade São Judas)
 Mestre em Eng. Elétrica (Centro Universitário da FEI)
 Doutor em Ciências (POLI – USP)

Atividade profissional:

- Professor Associado na FATEC- Santo André (Eletrônica Automotiva)
- Chefe do Departamento de Projetos de Máquinas na Mahle Anéis

ekitani@lsi.usp.br www.lsi.usp.br/~edson

201

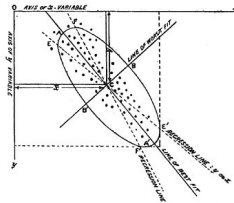
201

HISTÓRICO DO PCA

Descrito inicialmente por Karl Pearson no artigo, “*On lines and planes of closest fit to systems of points in space*”, *Philosophical Magazine*, 1901.



1857 - 1936



“In many physical, statistical, and biological investigations it is desirable to represent a system of points in plane, three or higher dimensional space by the best-fitting straight line or plane” (Pearson, 1901).

Prof. Edson Kitani - 04/2015

202

202

203

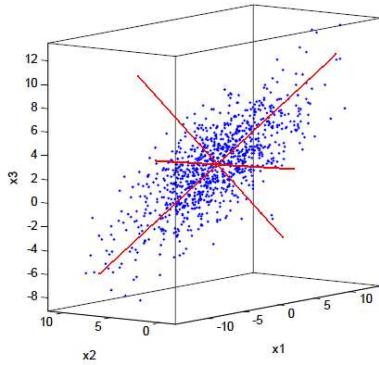
Aula 12
(de 15)
do Prof. Emilio
de PSI3471 de
03-abr-2019
se encerrou aqui

© Prof. Emilio Del Moral Hernandez

203

203

Para 3 Dimensões



$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i$$

$$\Sigma_{\mathbf{x}} = \frac{1}{(N-1)} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$$

$$\bar{\mathbf{x}} = [0 \quad 5 \quad 2]^T$$

$$\Sigma_{\mathbf{x}} = \begin{bmatrix} 25 & -1 & 7 \\ -1 & 4 & -4 \\ 7 & -4 & 10 \end{bmatrix}$$

Extraído de Osuna 2002

Prof. Edson Kitani - 04/2015

204

204

Exemplo Numérico com Matlab

```
x = [1, 2; 3, 3; 3, 3; 5, 5; 5, 4; 5, 6; 6, 5; 8, 7; 9, 8]
```

```
figure(1)
```

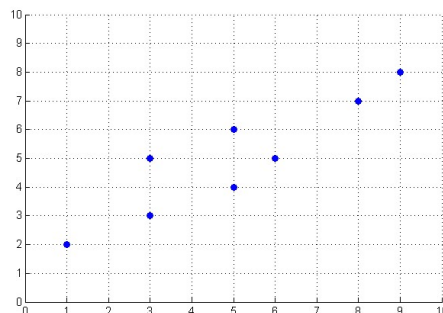
```
scatter(x(:,1),x(:,2))
```

```
Media = mean(x)
```

$$\bar{\mathbf{x}} = [5 \quad 5]^T$$

```
Sigma = cov(x,1)Nota
```

$$\Sigma_{\mathbf{x}} = \begin{bmatrix} 6,25 & 4,25 \\ 4,25 & 3,5 \end{bmatrix}$$



A normalização é realizado para N.

Prof. Edson Kitani - 04/2015

205

205

Autovalores (λ) e Autovetores (v)

Associado a cada matriz quadrada $A = ((a_{ij}))$ de ordem n , temos uma função

$$f(\lambda) = |A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix}$$

chamada de Função Característica da matriz A . E a função

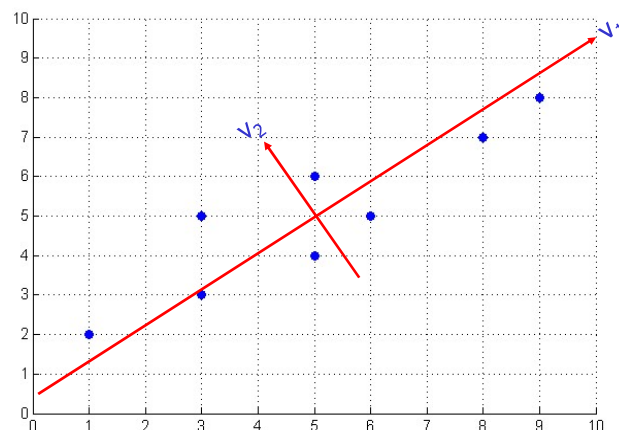
$$f(\lambda) = |A - \lambda I| = 0,$$

que pode ser expressa na forma polinomial

$$c_0 \lambda^n + c_1 \lambda^{n-1} + \dots + c_{n-1} \lambda + c_n = 0$$

é chamada de Equação Característica da matriz A .

Exemplo Numérico



Ortonormalização dos autovetores

$$\Phi_{pca} = \frac{\mathbf{x}^T \Phi(\Lambda)^{-\frac{1}{2}}}{(\sqrt{N-1})}$$

