Classificação Supervisionada para Dados de Elevada Dimensão

A. PEDRO DUARTE SILVA

Faculdade de Economia e Gestão /
Centro de Estudos em Gestão e Economia

Universidade Católica Portuguesa
Centro Regional do Porto

Classificação em Grandes Dimensões

Indice

1. Regras clássicas de classificação
2. Regras “diagonais” de classificação
3. Selecção de variáveis em dimensões elevadas
4. Como incorporar correlações quando p > n ?
5. Alguns resultados empíricos
6. Conclusões e perspectivas

Classificação em Grandes Dimensões

Formulação do Problema:

$(Y ; X)$ $\quad Y \in \{1,\ldots,k\}$ $\quad X \in \Re^p$

Pretende-se determinar uma regra capaz de prever Y dado X

Regra de Bayes: $\hat{Y} = \mathbf{argmax}_g \ \pi_g f_g(X)$

Pressupondo $X \mid Y \sim N_p(\mu_{(Y)}, \Sigma)$

$\Rightarrow$ Regra de Bayes:

$$\hat{Y} = \mathbf{argmin}_g \left(\frac{1}{2}(X-\mu_{(g)})^T \Sigma^{-1}(X-\mu_{(g)}) - \log \pi_g \right) =$$

$$= \mathbf{argmax}_g \left(\mu_{(g)}^T \Sigma^{-1} X - \frac{1}{2}\left(\mu_{(g)}^T \Sigma^{-1} \mu_{(g)} - \log \pi_g\right)\right)$$

Classificação em Grandes Dimensões

Regra de Fisher:

Dada uma amostra de observações com origem conhecida

$(Y_i ; X_i)$ $\quad i = 1,2, \ldots , n \quad \mathbf{n} = \sum_{g=1}^{k} \mathbf{n_g}$

$$\hat{\mathbf{Y}} = \mathbf{argmax_g} \left(\hat{\boldsymbol{\mu}}_{(g)}^{T} \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{X} - \frac{1}{2} (\hat{\boldsymbol{\mu}}_{(g)}^{T} \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\mu}}_{(g)} - \log \hat{\pi}_g \right)$$

$$= \mathbf{argmax_g} \left(\overline{\mathbf{X}}_{(g)}^{T} \mathbf{S}^{-1} \mathbf{X} - \frac{1}{2} \left(\overline{\mathbf{X}}_{(g)}^{T} \mathbf{S}^{-1} \overline{\mathbf{X}}_{(g)}^{T} - \log \frac{n_g}{n} \right) \right)$$

$$\overline{\mathbf{X}}_g = \frac{1}{\mathbf{n_g}} \sum_{Y_i = g} \mathbf{X_i} \qquad \mathbf{S} = \frac{\sum_{g=1}^{k} \sum_{Y_i = g} (\mathbf{X_i} - \overline{\mathbf{X}}_{(g)})(\mathbf{X_i} - \overline{\mathbf{X}}_{(g)})^{T}}{\mathbf{n - k}}$$

Classificação em Grandes Dimensões

Limitações da regra de Fisher:

* **Pressupostos restritivos**

  (modelo Gaussiano homocedástico)

* **Estimação ineficiente de $\mu_{(g)}$ e Σ^{-1}**

  (quando n não é suficientemente superior a p)

  - **quando p > n, S é singular**
  - **quando n > p, mas p e n estão próximos, S^{-1} é instável**
  - **quando p >> n, os erros acumulados na estimação de $\mu_{(g)}$ assumem um peso excessivo**
  - **frequentemente a maioria das variáveis originais apenas acrescenta ruído**

Classificação em Grandes Dimensões

Análise Discriminante Diagonal

Naive Bayes

$$\hat{\mu}_{(g)} = \overline{X}_{(g)} \qquad \hat{\pi}_g = \frac{n_g}{n} \qquad \hat{\Sigma} = \hat{D} = \text{diag}(S)$$

Nearest Shruken Centroids

Tibshirani, Hastie, Narasimhan e Chu (2003)

$$\hat{\mu}^*_{(g)} = \hat{\mu} + \sqrt{\frac{1}{n_g} - \frac{1}{n}}\ \hat{D}^{0.5}\ d^*_g$$

$$d^*_g(j) = \text{sign}(d_g(j))\ (\max(d_g(j) - t^*; 0)$$

$$d_g = \frac{\hat{\mu}_{(g)} - \hat{\mu}}{\theta_g \sqrt{\frac{1}{n_g} - \frac{1}{n}}}\ \hat{D}^{-0.5} \qquad \hat{\mu} = \frac{1}{n} \sum_{g=1}^{k} n_k\ \hat{\mu}_{(g)}$$

$t^*, \theta_1, \ldots, \theta_k$ obtidos por validação cruzada

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um critério de ordenação

Problemas com 2 grupos

Valor das estatísticas t para a comparação das médias por grupo

Problemas com mais de 2 grupos

* Valor das estatísticas F para uma ANOVA a um factor
* Diferenças (soma das ?) estandartizadas entre as médias de grupo e as média globais
* Menor diferenças estandartizada entre todos os pares de grupos

Defenir um limiar de selecção

Pontos criticos de distribuições nulas ?

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um limiar de selecção

"The classical frequentist theory of hypotesis testing ... has a claim to being the twentieth century's most influential piece of applied mathematics."

"… something new is happening in the twenty-first century "

Bradly Efron

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um limiar de selecção

"Something new is happening in the twenty-first century... "

- **bases de dados de dimensão massiva**
- **centenas (ou milhares) de testes são realizados em simultaneo**
- **métodos clássicos para controlar probabilidades de erro são claramente inadequados**
  - **métodos que controlam probabilidades individuais de erro são demasiado liberais**
  - **métodos que controlam probabilidades globais de erro são demasiado conservadores**

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um limiar de selecção

Controlo da taxa de falsas descobertas **(Benjamini e Hochberg 1995)**

Dada uma sucessão de p testes onde hipoteses nulas rejeitadas correspondem a "descobertas" de efeitos "interessantes"

Frequentemente, **em vez de**:

- **garantir que a probabilidade de cada falsa descoberta é inferior a α**

ou:

- **garantir que a probabilidade de haver alguma falsa descoberta é inferior a α**

É mais relevante:

- **garantir que a proporção esperada de falsas descobertas é inferior a α**

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um limiar de selecção

Controlo da taxa de falsas descobertas

Dada uma sucessão de p testes independentes, com valores de prova ordenados $\pi_1, \ldots, \pi_p$

Rejeitam-se as hipoteses nulas (H_{0j}) em que $j \leq k$, com

$$k = \max\left\{ j : \pi_j \leq \frac{j}{p}\,\alpha \right\}$$

(Benjamini e Hochberg 1995)

Dada uma sucessão de p testes eventualmente dependentes, com valores de prova ordenados $\pi_1, \ldots, \pi_p$

Rejeitam-se as hipoteses nulas (H_{0j}) em que $j \leq k$, com

$$k = \max\left\{ j : \pi_j \leq \frac{j}{p\sum_{i=1}^{p}\frac{1}{i}}\,\alpha \right\}$$

(Benjamini e Yekutieli 2001)

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um limiar de selecção

Extensões: taxas locais de falsas descobertas

Dada uma sucessão de p testes e estatísticas ordenadas, z_1, ..., z_p com

$P_0 = P(H_0)$; $P_1 = P(H_1)$ $\quad$ $f_0(z)$; $f_1(z)$; $f(z) = P_0\, f_0(z) + P_1\, f_1(z)$

Taxa local de falsas descobertas: **$fdr(z) = P_0\, f_0(z) / f(z)$**

Taxa local de falsas não-descobertas: **$fndr(z) = 1 - fdr(z)$**

Extensões: taxas baseados em "distribuições nulas empiricas"

Estimar $f_0(z)$ com base nos menores $\hat{p}_0$ valores de z_1, ..., z_p (Efron 2004)

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um limiar de selecção

Higher Criticism **(Donoho e Jin 2004)**

Dada uma sucessão de p testes independentes, com valores de prova ordenados $\pi_1, \ldots, \pi_p$

$$HC(j;\pi_j) = \sqrt{p}\,\frac{(j/p) - \pi_j}{\sqrt{(j/p)(1-(j/p))}}$$

mede a evidência contra H_{0j} baseada na sucessão completa dos p testes

Se existirem "efeitos" (violações de hipoteses nulas) de pequena magnitude numa percentagem reduzida de testes,

$$HC^* = \max_{j \le \alpha_0} HC(j;\pi_j)$$

tem propriedades adaptativas, assintóticamente óptimas

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um limiar de selecção

Higher Criticism

Num modelo gaussiano homecedástico de classificação entre dois grupos com:

- Variáveis independentes
- Efeitos (diferenças entre as médias dos grupos) raros
- Efeitos fracos
- Variáveis ordenadas com base em estatisticas t

quando p $\to\infty$

HC* é assintóticamente equivalente ao limiar óptimo de selecção

(Donoho e Jin 2009)

Classificação em Grandes Dimensões

Métodos modernos de selecção de variáveis

Defenir um limiar de selecção

Uma proposta para problemas com efeitos raros e maoritáriamente fracos

- **1** Incluir todas as variáveis que satisfaçam o critério de Benjamini e Yekutieli

- **2** Estimar uma "distribuição nula empirica" e calcular os respectivos valores de prova

- **3** Escolher o limiar HC* baseado nos valores de prova calculados no passo anterior

Classificação em Grandes Dimensões

Como incorporar correlações ?

Thomaz e Gilles's "Novas FDL"

$$\mathbf{S} = \sum_{m=1}^{p} \lambda_m \mathbf{v}_m \mathbf{v}_m^T \qquad \mathbf{S}^{-1} = \sum_{m=1}^{p} \frac{1}{\lambda_m} \mathbf{v}_m \mathbf{v}_m^T \qquad \hat{\Sigma}^{-1} = \sum_{m=1}^{p} \frac{1}{\max(\lambda_m, \bar{\lambda})} \mathbf{v}_m \mathbf{v}_m^T$$

Estimadores "encolhidos" e regularizados

$$\hat{\Sigma} = \rho_1 I_p + \rho_2 S$$

Guo, Hastie e Tibshirani (2007)

ρ_1 e ρ_2 obtidos por validação-cruzada

Xu, Brock e Parrish (2009)

James-Stein formulas para ρ_1 e $\rho_2 = 1 - \rho_1$

ou:

$$\Sigma = \mathbf{V}^{1/2} \mathbf{R} \mathbf{V}^{1/2}$$

$$\hat{\mathbf{R}}^* = (1 - \rho_1)\hat{\mathbf{R}}$$

$$\hat{\mathbf{V}}^*(j, j) = \rho_2 \, me_j(\hat{\mathbf{V}}(j, j)) + (1 - \rho_2)\hat{\mathbf{V}}(j, j)$$

Ahdesmaki e Strimmer (2009)

$$\hat{\pi}^*_g = \rho_3 \frac{1}{k} + (1 - \rho_3)\frac{n_g}{n}$$

James-Stein formulas para ρ_1, ρ_2 e ρ_3

Classificação em Grandes Dimensões

Como incorporar correlações ?

Uma proposta baseada em modelos factoriais

$$X_i = \mu_{(Yi)} + B\ f_i + \varepsilon_i \qquad f_i \in \Re^q \qquad \varepsilon_i \in \Re^P \qquad q << p$$

$$f_i \sim N_q(0, I_q) \qquad \varepsilon_i \sim N_p(0, D_\varepsilon) \qquad \forall j\ D_\varepsilon(j) > k_0 \in \Re_0$$

$\Rightarrow$

$$\Sigma = B\ B^T + D_\varepsilon$$

$$\Sigma^{-1} = D_\varepsilon^{-1} - D_\varepsilon^{-1} B\ [I_q + B^T D_\varepsilon^{-1} B]^{-1} B^T D_\varepsilon^{-1}$$

$$\hat{\Sigma}_{Fctq} = \hat{B}\hat{B}^T + \hat{D}_\varepsilon \quad ; \quad \hat{B}, \hat{D}_\varepsilon = \arg\min_{\hat{B},\hat{D}_\varepsilon} \| \hat{V}^{-1/2}\hat{\Sigma}_{Fctq}\hat{V}^{-1/2} - \hat{V}^{-1/2} S\, \hat{V}^{-1/2} \|^2$$

Classificação em Grandes Dimensões

Singh's Prostate Cancer Data – p=6033; n=50+52

Rule	Error Estimate (std error)	# Variables kept (min – median - max)
Fisher's LDA*	0.2146 (0.0101)	58 – **134.5** – 421
Naive Bayes*	**0.0670** (0.0052)	58 – **134.5** – 421
Support Vector Machines*	**0.0642** (0.0052)	58 – **134.5** – 421
Nearest Shruken Centroids	0.0838 (0.0063)	108 – **356** – 1771
Regularized DA	**0.0741** (0.0053)	82 – **390** – 1201
Shrunken DA*	**0.0650** (0.0051)	58 – **134.5** – 421
Factor-based LDA* (q=1)	**0.0641** (0.0052)	58 – **134.5** – 421
NLDA*	**0.0720** (0.0052)	58 – **134.5** – 421

* After variable selection by the maximum of FDR (False Discovery Rates) and HC (Higher Criticism), both derived from Independence based T-scores. The p-values used in the HC computations are derived from empirical Null distributions

Classificação em Grandes Dimensões

Golubs's Leukemia Data -- p = 7 129 ; n = 47+25

Rule	Error Estimate (std error)	# Variables kept (min – median - max)
Fisher's LDA*	0.2558 (0.0109)	326 – **478** – 712
Naive Bayes*	0.480 (0.0085)	326 – **478** – 712
Support Vector Machines*	0.0405 (0.0049)	326 – **478** – 712
Nearest Shruken Centroids	**0.0201** (0.0039)	703 – **3166** – 7129
Regularized DA	0.0491 (0.0062)	12 – **1934** – 7124
Shrunken DA*	0.0276 (0.0044)	326 – **478** – 712
Factor-based LDA* (q=1)	**0.0174** (0.0034)	326 – **478** – 712
NLDA*	0.1510 (0.0085)	326 – **478** – 712

* After variable selection by the maximum of FDR (False Discovery Rates) and HC (Higher Criticism), both derived from Independence based T-scores. The p-values used in the HC computations are derived from empirical Null distributions

Classificação em Grandes Dimensões

Golubs's Leukemia Data -- p = 7 129 ; n = 47+25

Rule	Error Estimate (std error)	# Variables kept (min – median - max)
Fisher's LDA*	0.2558 (0.0109)	326 – **478** – 712
Naive Bayes*	0.480 (0.0085)	326 – **478** – 712
Support Vector Machines*	0.0405 (0.0049)	326 – **478** – 712
Nearest Shruken Centroids	**0.0201** (0.0039)	703 – **3166** – 7129
Regularized DA	0.0491 (0.0062)	12 – **1934** – 7124
Shrunken DA*	0.0276 (0.0044)	326 – **478** – 712
Factor-based LDA* (q=1)	**0.0174** (0.0034)	326 – **478** – 712
NLDA*	0.1510 (0.0085)	326 – **478** – 712

* After variable selection by the maximum of FDR (False Discovery Rates) and HC (Higher Criticism), both derived from Independence based T-scores. The p-values used in the HC computations are derived from empirical Null distributions

Classificação em Grandes Dimensões

Alon's Colon Data -- p = 2 000 ; n = 40+22

Rule	Error Estimate (std error)	# Variables kept (min – median - max)
Fisher's LDA*	0.3285 (0.0143)	3 – **71.5** – 200
Naive Bayes*	0.2275 (0.0133)	3 – **71.5** – 200
Support Vector Machines*	0.1576 (0.0095)	3 – **71.5** – 200
Nearest Shruken Centroids	**0.1563** (0.0098)	7 – **39** – 527
Regularized DA	0.2174 (0.0126)	14 – **425** – 2000
Shrunken DA*	0.1865 (0.0100)	3 – **71.5** – 200
Factor-based LDA* (q=1)	**0.1746** (0.0098)	3 – **71.5** – 200
NLDA*	0.2614 (0.0114)	3 – **71.5** – 200

* After variable selection by the maximum of FDR (False Discovery Rates) and HC (Higher Criticism), both derived from Independence based T-scores. The p-values used in the HC computations are derived from empirical Null distributions

Classificação em Grandes Dimensões

Simulation Experiment -- Independent Data

p = 5 000 ; n=50 ; π_0 = 0.5 ; noise level = 97% ; Mah dist = 6

Rule	**Error Estimate (std error)**	**# Variables kept (min – median - max)**
Fisher's LDA*	0.0898 (0.0114)	6 – **16** – 478
Naive Bayes*	0.0303 (0.0046)	6 – **16** – 478
Support Vector Machines*	0.0681 (0.0097)	6 – **16** – 478
Nearest Shruken Centroids	**0.0148** (0.0013)	7 – **39** – 527
Regularized DA	0.0323 (0.0028)	6 – **253.5** – 1452
Shrunken DA*	0.0228 (0.0039)	6 – **16** – 478
Factor-based LDA* (q=1)	0.0260 (0.0036)	7 – **16** – 351
NLDA*	0.0284 (0.0044)	6 – **16** – 478

* After variable selection by the maximum of FDR (False Discovery Rates) and HC (Higher Criticism), both derived from Independence based T-scores. The p-values used in the HC computations are derived from empirical Null distributions

Classificação em Grandes Dimensões

Simulation Experiment -- Correlated Data

p = 5 000 ; n=50 ; π_0 = 0.5 ; noise level = 97% ; Mah dist = 6

Rule	Error Estimate (std error)	# Variables kept (min – median - max)
Fisher's LDA*	0.1668 (0.0114)	1 – **8** – 801
Naive Bayes*	0.1815 (0.0046)	1 – **8** – 801
Support Vector Machines*	0.1872 (0.0097)	1 – **8** – 801
Nearest Shruken Centroids	0.1637 (0.0043)	3 – **11** – 104
Regularized DA	**0.1553** (0.0069)	1 – **6** – 733
Shrunken DA*	0.1882 (0.0048)	4 – **13** – 477
Factor-based LDA* (q=1)	**0.1511** (0.0031)	1 – **7** – 507
NLDA*	0.1987 (0.0061)	4 – **13** – 477

* After variable selection by the maximum of FDR (False Discovery Rates) and HC (Higher Criticism), both derived from Independence based T-scores. The p-values used in the HC computations are derived from empirical Null distributions

Classificação em Grandes Dimensões

Simulation Experiment -- Guo, Hastie and Tibshirani (2007) setup

p=10 000; n=100+100 ; 100 Independent Blocks ; ρ = 0.90

Rule	Selection Criterion*	Error Estimate	2 std errors
Shrunken DA	I_HC	**0.0029**	0.0006
Factor-based LDA (q=1)	I_HC	0.0122	0.0020
NLDA	I_HC	0.0055	0.0008
Nearest Shruken Centroids	-------	0.0149	0.0014
Naive Bayes	I_HC	0.0136	0.0019
Support Vector Machines	I_HC	0.0052	0.0008
Fisher's LDA	I_HC	0.0435	0.0133

* I - Independence based T-scores
HC - Higher Criticism

Classificação em Grandes Dimensões

Simulation Experiment -- Guo, Hastie and Tibshirani (2007) setup

p=10 000; n=100+100 ; 10 Independent Blocks ; ρ = 0.90

Rule	Selection Criterion*	Error Estimate	2 std errors
Shrunken DA	I_HC	0.4973	0.0006
Factor-based LDA (q=1)	I_HC	0.4955	0.0008
NLDA	I_HC	**0.0186**	0.0037
Nearest Shruken Centroids	-------	0.4951	0.0009
Naive Bayes	I_HC	0.4942	0.0011
Support Vector Machines	I_HC	0.4963	0.0006
Fisher's LDA	I_HC	0.0293	0.0131

* I - Independence based T-scores
  HC - Higher Criticism

Classificação em Grandes Dimensões

Conclusões

Escolha do numero de variáveis

- A combinação do controlo da taxa de falsas descobertas com a "Higher Criticism" estatística, parece dar bons resultados

Utilidade das correlações empiricas

Regras bem-condicionadas de classificação que incorporam correlações empiricas, frequentemente tem:

desempenho semelhante a regras diagonais para dados independentes

desempenho (ligeiramente) superior a regras diagonais para dados correlacionados

Classificação em Grandes Dimensões

Conclusões

Utilidade das correlações empiricas

Nenhuma das abordagens alternativas para a incorporação de correlações leva a resultados uniformemente superiores às restantes

Perspectivas e questões em aberto

Dever-se-ão incorporar correlações tambem na selecção de variáveis?

Quando e como?

Correlation Adjusted t-scores ? Zuber e Strimmer (2009)

Combinar "testes de informação adicional" com controlo da taxa de erro e a estatistica HC ?

Que estimador de Σ^{-1} utilizar ?

Classificação em Grandes Dimensões

Perspectivas e questões em aberto

Qual o melhor critério de ordenação para problemas com mais de dois grupos?

Quais as propriedades assintóticas das regras de classificação quando $p \to \infty$?

Qual a relevância dessas propriedades ?

Classificação em Grandes Dimensões

Referências

Ahdesmaki, P. and Strimmer, K. (2009). Feature selection in "omics"prediction problems using cat scores and non-discovery rate control. *rXiv,stat.AP:0903.2003v1.*

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society* B 57, 289-300.

Benjamini, Y. and Yekutileli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* 29, 1165-1188.

Bickel, P.J. and Levina, E. (2004). Some theory for Fisher's linear discriminant function, naive Bayes and some alternatives when there are many more variables than observations . *Bernoulli* 10, 989-1010.

Donoho, D. and Jin, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Annals of Statistics*, 32, 962-944.

Donoho, D. and Jin, J. (2008). Higher criticism thresholding: Optimal feature selection when useful features are rare and weak. *Proc. Natl. Acad. Sci*, USA 105, 14790-14795.

Donoho, D. and Jin, J. (2009). Feature selection by higher criticism thresholding: Optimal phase diagram. *Philosophical Transactions of the Royal Society A,* 367, 4449-4470.

Duarte Silva, A.P. (2009). Linear Discriminant Analysis with more Variables than Observations. A not so Naïve Approach. To appear In: *Classification as a Tool for Research. Proceedings of the 11th IFCS Biennial Conference and 33rd Annual Conference of the Gesellschaft für Klassifikation*. Dresden, Germany.

Duarte Silva, A.P. High-Dimensional Classication in the Presence of Correlation: A Factor Model Approach (submetido para publicação).

Classificação em Grandes Dimensões

Referências (cont.)

Efrom, B. (2008). Microarrays, empirical Bayes and the two-groups model. *Statistical Science* 1, 1-22.

Guo, Y., Hastie, T. and Tibshirani, T. (2007). Regularized discriminant analysis and its application in microarrays. *Biostatistics* 8, 86-100.

Strimmer, K. (2008). A unified approach to false discovery rate estimation. *BMC Bioinformatics*, 9, 303.

Tibshirani, R., Hastie, B., Narismhan, B. and Chu, G. (2003). Class prediction by nearest shrunken centroids with applications to DNA microarrays. *Statistical Science*, 18, 104-117.

Thomaz, C.E. and Gillies, D.F. (2005). A maximum uncertainty lda-based approach for limited sample size problems with application to face recognition. In*: 18th Brazilian Symposium on computer Graphics and Image Processimg. SIBGRAPI 2005*, 89-96.

Xu, P., Brock, G.N., and Parrish, R. (2009). Modified linear discriminant analysis approaches for classification of high-dimensional microarray data. *Computational Statistics and Data Analysis*, 53, 1674-1687.

Zu, V. and Strimmer, K. (2009). Gene ranking and biomaker discovery and correlation.

rXiv,stat.AP:0902.0751v1.

Classificação em Grandes Dimensões

Obrigado pela vossa atenção !!

Questões ??