

Philippe Bastien

L'Oreal Research, Clichy, France (pbastien@rd.loreal.com)

AMRI

L'ORÉAL
Research & Innovation

Régression et multicollinéarité

Régression et multicollinéarité

$$= \beta + \varepsilon$$

$$\beta = \frac{1}{\|X\|} \|y - \beta\|$$

$$\rightarrow \beta = \frac{1}{\|X\|} \|y - \beta\|$$

$$\beta = \sigma \frac{1}{\|X\|}$$

$$= \frac{1}{\|X\|}$$

$$= \frac{1}{\|X\|} \rightarrow \infty \rightarrow$$

$$\beta \rightarrow \infty$$

➤ 1) Diagonaliser

$$= \Sigma' = \sum \sqrt{\lambda}$$

$$\boxed{} = \sqrt{\lambda} \boxed{} + \sqrt{\lambda} \boxed{}$$

$$' - = \Sigma^{-} '$$

$$\beta = ' - ' = \Sigma^{-} ' \Sigma ' = \Sigma^{-} '$$



2) *Tridiagonaliser* $X'X$

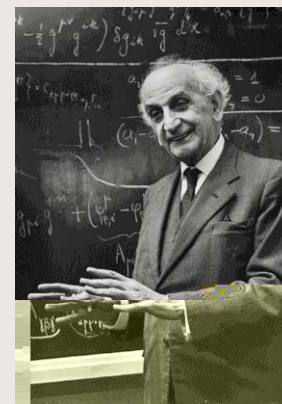
$$= ' = '$$

$$= \quad , \quad = \quad , \quad , \quad , \quad , \quad - \quad ,$$

$$\beta = (\quad ' \quad ' \quad)^{-1} \quad ' \quad ' \quad '$$

$$\beta = (\quad ' \quad)^{-1} \quad ' \quad \beta$$

β



Cornelius Lanczos (1893-1974)
mathématicien et physicien hongrois

3) Une factorisation particulière de X

$$= \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} = \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} =$$

$$\beta = \begin{pmatrix} - & - \\ - & - \end{pmatrix} = \begin{pmatrix} - & - \\ - & - \end{pmatrix} =$$

$$=$$

$$= \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \parallel \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$$

$$= \begin{pmatrix} - & - \\ - & - \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \cong$$

$$+ \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} - & - \\ - & - \end{pmatrix} \cong$$

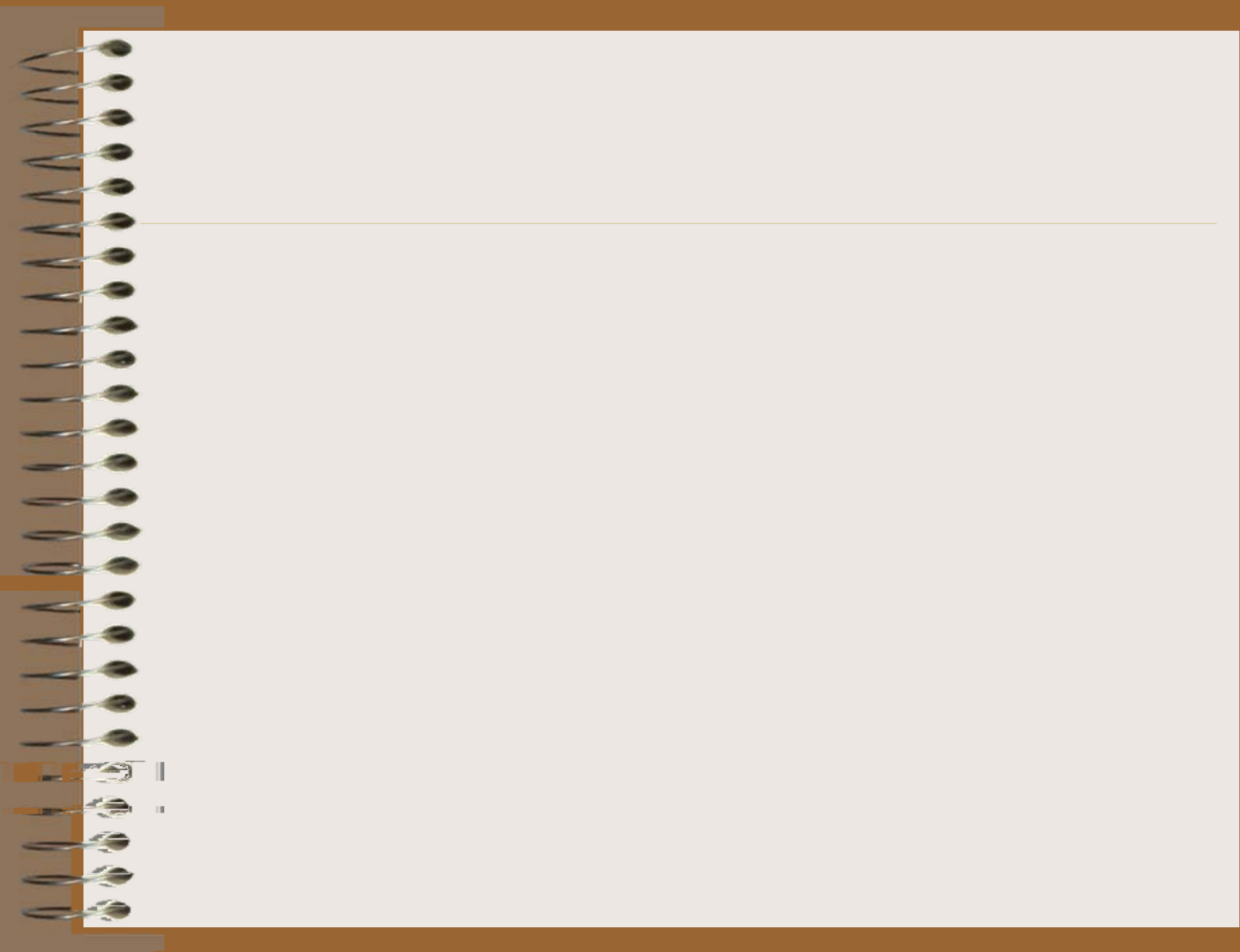
$$=$$

$$=$$

Exemple: Régression et multicollinéarité

Obs	Obs	AUTO	CYL	PUI	LON	LAR	POIDS	VITESSE	PRIX
1	1	ALFASUD-TI-1350	1350	79	393	161	870	165	30570
2	2	AUDI-100-L	1588	85	468	177	1110	160	39990
3	3	SIMCA-1307-GLS	1294	68	424	168	1050	152	29600
4	4	CITROEN-GS-CLUB	1222	59	412	161	930	151	28250
5	5	FIAT-132-1600GLS	1585	98	439	164	1105	165	34900
6	6	LANCIA-BETA-1300	1297	82	429	169	1080	160	35480
7	7	PEUGEOT-504	1796	79	449	169	1160	154	32300
8	8	RENAULT-16-TI	1565	55	424	163	1010	140	32000
9	9	RENAULT-30-TS	2664	128	452	173	1320	180	47700
10	10	TOYOTA-COROLLA	1166	55	399	157	815	140	26540
11	11	ALFETTA-1-66	1570	109	428	162	1060	175	42395
12	12	PRINCESS-1800-HI	1798	82	445	172	1160	158	33990
13	13	DATSUN-2001	1998	115	469	169	1370	160	43980
14	14	TAUNUS-2000-GL	1993	98	438	170	1080	167	35010
15	15	RANCHO	1442	80	431	166	1129	144	39450
16	16	MAZDA-9291	1769	83	440	165	1095	165	27900
17	17	OPEL-RECORD-L	1579	100	450	173	1120	173	32700
18	18	LADA-1500	1294	68	404	161	955	140	22100

Exemple: Régression et multicolinéarité



Exemple: Régression et multicolinéarité : PLSR

The PLS Procedure

Variation en pourcentage expliquée
par Partial Least Squares Factors

Nombre de facteurs extraits	Effets du modèle		Variables dépendantes	
	Actuel	Total	Actuel	Total
1	73.6230	73.6230	60.8374	60.8374

Parameter Estimates for Centered and Scaled Data

PRIX

Intercept	0.0000000000
CYL	0.1457852413
PUI	0.1823397520
LON	0.1469668392
LAR	0.1247976334
POIDS	0.1719738622
VITESSE	0.1328131564

Parameter Estimates

PRIX

Intercept	-39940.36629
CYL	2.56208
PUI	58.80660
LON	43.68699
LAR	154.34048
POIDS	8.25174
VITESSE	71.89164

Exemple: Régression et multicolinéarité

Obs	Obs	AUTO	CYL	PUI	LON	LAR	POIDS	VITESSE	PRIX	Pred
1	1	ALFASUD-TI-1350	1350	79	393	161	870	165	30570	29223
2	2	AUDI-100-L	1588	85	468	177	1110	160	39990	37553
3	3	SIMCA-1307-GLS	1294	68	424	168	1050	152	29600	31418
4	4	CITROEN-GS-CLUB	1222	59	412	161	930	151	28250	28038
5	5	FIAT-132-1600GLS	1585	98	439	164	1105	165	34900	35354
6	6	LANCIA-BETA-1300	1297	82	429	169	1080	160	35480	33445
7	7	PEUGEOT-504	1796	79	449	169	1160	154	32300	35649
8	8	RENAULT-16-TL	1565	55	424	163	1010	140	32000	29384
9	9	RENAULT-30-TS	2664	128	452	173	1320	180	47700	44692
10	10	TOYOTA-COROLLA	1166	55	399	157	815	140	26540	24734
11	11	ALFETTA-1-66	1570	109	428	162	1060	175	42395	35521
12	12	PRINCESS-1800-HL	1798	82	445	172	1160	158	33990	36407
13	13	DATSUN-200L	1998	115	469	169	1370	160	43980	41322
14	14	TAUNUS-2000-GL	1993	98	438	170	1080	167	35010	37219
15	15	RANCHO	1442	80	431	166	1129	144	39450	32577
16	16	MAZDA-9295	1769	83	440	165	1095	165	27900	35059
17	17	OPEL-RKORD-L	1979	100	459	173	1120	173	32700	39443
18	18	LADA-1300	1294	68	404	161	955	140	22100	27817

Exemple: Régression et multicollinéarité: PLSR

The PLS Procedure

Variation en pourcentage expliquée
par Partial Least Squares Factors

Nombre de facteurs extraits	Effets du modèle		Variables dépendantes	
	Actuel	Total	Actuel	Total
1	73.6230	73.6230	60.8374	60.8374
2	9.0379	82.6609	6.2420	67.0794
3	9.2571	91.9181	2.3752	69.4546
4	4.6704	96.5885	0.9020	70.3566
5	2.4807	99.0691	0.4523	70.8090
6	0.9309	100.0000	0.1013	70.9103

Parameter Estimates for Centered and Scaled Data

PRIX

Intercept	0.0000000000
CYL	-.1994486089
PUI	0.8749117597
LON	-.0505879950
LAR	0.1687469748
POIDS	0.2620680197
VITESSE	-.2052720153

Parameter Estimates

PRIX

Intercept	-8239.362677
CYL	-3.505182
PUI	282.168803
LON	-15.037660
LAR	208.693773
POIDS	12.574678
VITESSE	-111.113551

Exemple: Régression et données manquantes

Obs	Obs	AUTO	CYL	PUI	LON	LAR	POIDS	VITESSE	PRIX
1	1	ALFASUD-TI-1350	.	79	393	161	870	165	30570
2	2	AUDI-100-L	1588	.	468	177	1110	160	39990
3	3	SIMCA-1307-GLS	1294	68	.	168	1050	152	29600
4	4	CITROEN-GS-CLUB	1222	59	412	.	930	151	28250
5	5	FIAT-132-1600GLS	1585	98	439	164	.	165	34900
6	6	LANCIA-BETA-1300	1297	82	429	169	1080	.	35480
7	7	PEUGEOT-504	1796	79	449	169	.	154	32300
8	8	RENAULT-16-TL	1565	55	424	.	1010	140	32000
9	9	RENAULT-30-TS	2664	128	.	173	1320	180	47700
10	10	TOYOTA-COROLLA	1166	.	399	157	815	140	26540
11	11	ALFETTA-1-66	.	109	428	162	1060	175	42395
12	12	PRINCESS-1800-HL	1798	.	445	172	1160	158	33990
13	13	DATSUN-200L	1998	115	.	169	1370	160	43980
14	14	TAUNUS-2000-GL	1993	98	438	.	1080	167	35010
15	15	RANCHO	1442	80	431	166	.	144	39450
16	16	MAZDA-9295	1769	83	440	165	1095	.	27900
17	17	OPEL-RKORD-L	1979	100	459	173	.	173	32700
18	18	LADA-1300	1294	68	404	.	955	140	22100

```
228 proc pls data=lib.voitureswithmissing METHOD=PLS (ALGORITHM=NIPALS)CV=BLOCK details;
229     model PRIX=CYL PUI LON LAR POIDS VITESSE/solution;
230 run;
```

ERROR: No valid observations.

NOTE: 18 observation(s) lue(s) dans la table LIB.VOITURESWITHMISSING.

NOTE: Procédure PLS a utilisé (Durée totale du traitement) :

temps réel 0.00 secondes
temps UC 0.00 secondes

Exemple: Régression et multicollinéarité

PLS-NIPALS

Var ID

Coeff (PRIX)

constant

CYL

PUI

LON

LAR

POIDS

VITESSE

Régularisation

Nécessité de régularisation afin d'obtenir un modèle prédictif stable

Régularisation et biais



β



Compromis biais-variance

$$\begin{aligned} \beta &= (\beta - \beta)' (\beta - \beta) \\ &= (\beta - \beta)' (\beta - \beta) + (\beta - \beta)' (\beta - \beta) \\ &= \quad + \quad \beta \end{aligned}$$

Quel est le biais de PLS ?

Régression biaisée

$$\hat{E}_{OLS} = \sum_{i=1}^r \frac{u_i \varphi}{\sqrt{\rho}} v_i \quad \hat{E}_T$$

$$\hat{E}_{\text{biaisé}} = \sum_{i=1}^r f(\rho_i) \hat{O}_i \quad E$$

$f(\rho) = 1 - \rho$ diminution de la variance de la $i^{\text{ème}}$ composante

Régression ridge et biais

$$\beta = \underset{\beta}{\arg \min} \|\beta\|^2 + \lambda \|\beta\| \rightarrow \beta = (\text{'} + \lambda)^{-1} \text{'}$$

$$\beta = \sum_{\beta} \frac{\lambda}{\lambda + \lambda} \beta \quad \lambda < \quad \beta = \frac{\rho}{+\lambda} \quad \frac{\rho}{+\lambda} \quad \frac{\rho}{+\lambda} \quad \frac{\rho}{+\lambda}$$

Régression PLS et biais

$$\beta = \sum_{\mu} -\prod_{\mu} -\frac{\lambda}{\mu} \beta$$

$$\mu =$$

$$=$$

λ

{ ' ' ' ' ' }

||

| λ |

>

NIPALS (Wold H., 1966)



Méthode de la puissance itérée (Hotelling, 1936)

$$= \lambda$$

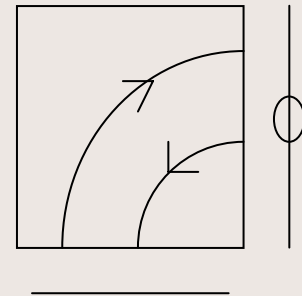
$$= \lambda$$

λ

$$\lambda = \sum \alpha \lambda \rightarrow \alpha \lambda \rightarrow \infty$$

$$\lambda > \lambda > \lambda$$

$$= \sum \alpha$$



Méthode de la puissance itérée (Hotelling, 1936)

$$\lambda' = \lambda$$

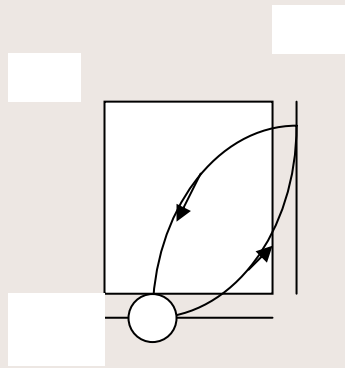
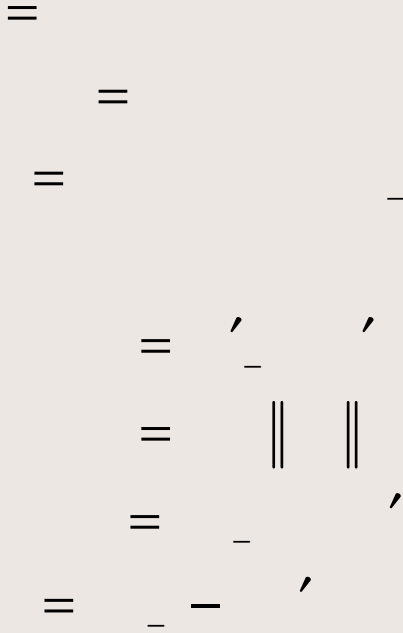
$$\lambda = \frac{\| \quad \|}{\| \quad \|} + \frac{\| \quad \|}{\| \quad \|}$$

$$\lambda' = \lambda' - \lambda$$

$$\lambda' = \lambda' - \sum_{i=1}^n \lambda_i$$

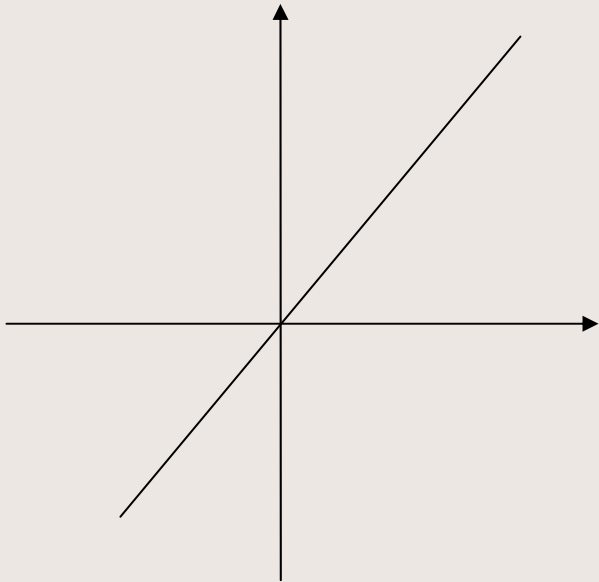
Méthode de la puissance itérée:

Algorithm NIPALS



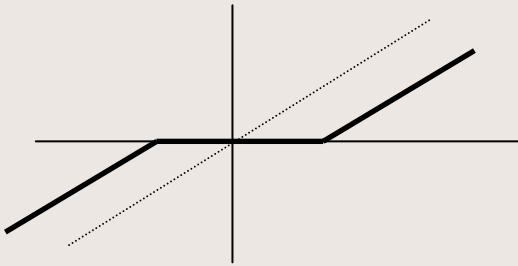
NIPALS et données manquantes

$$= \frac{-}{,}$$



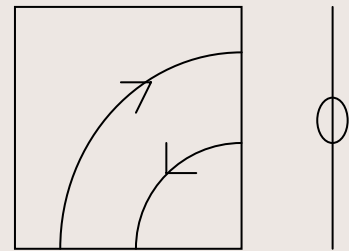
Sparse PCA

$$\lambda \quad (| \cdot | - \lambda)_+$$



Sparse PCA : Algorithmme

$$\begin{aligned}
 & \Sigma' = \\
 & = \\
 & = \delta \\
 & = \lambda \left(\begin{array}{c} ' \\ - \end{array} \right) \\
 & = - \parallel - \parallel \\
 & = \\
 & = \parallel \parallel \\
 & = - \delta '
 \end{aligned}$$

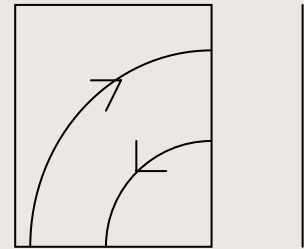


Algorithmme PLS1-NIPALS

$$\begin{aligned}
 &= \\
 &= \begin{matrix} ' \\ - \\ ' \end{matrix} \\
 \parallel & \parallel = \\
 &= \begin{matrix} ' \\ - \\ ' \end{matrix} \\
 &= \begin{matrix} ' \\ - \\ ' \end{matrix} \\
 &= \begin{matrix} ' \\ - \\ ' \end{matrix} \\
 &= \begin{matrix} ' \\ - \\ ' \end{matrix} \\
 &= \sum_{\underline{=}} \begin{matrix} ' \\ - \\ ' \end{matrix} = \sum_{\underline{=}} \begin{matrix} ' \\ - \\ ' \end{matrix} - = \sum_{\underline{=}} \begin{matrix} ' \\ - \\ ' \end{matrix} \\
 &= \begin{matrix} ' \\ - \\ ' \end{matrix} - \Pi(\begin{matrix} ' \\ - \\ ' \end{matrix}) \\
 &= \\
 &=
 \end{aligned}$$

Sparse PLS: Algorithm

$$\begin{aligned}
 & \mathbf{X}' = \Delta \mathbf{X}' \\
 & \mathbf{X}' = \mathbf{X}' \\
 & \mathbf{X}' = \mathbf{X}' \\
 & \mathbf{X}' = \mathbf{X}' \\
 & \mathbf{X}' = \lambda \mathbf{X}' \quad \| \quad \| = \\
 & \mathbf{X}' = \lambda \mathbf{X}' \quad \| \quad \| = \\
 & \mathbf{X}' = \mathbf{X}' \\
 \xi & = \mathbf{X}' \\
 & = \mathbf{X}' \\
 & = \mathbf{X}' \xi \quad \xi' \xi \\
 & = \mathbf{X}' \xi \quad \xi' \xi \\
 & = \mathbf{X}' \xi \\
 & = \mathbf{X}' \xi
 \end{aligned}$$



Régression PLS et très grande dimension

Kernel PLS

$\gg \rightarrow$

$\gg \rightarrow$

' '

' '

' = ' \rightarrow =

Régression PLS et très grande dimension

Canonical PLS

$$\beta = - (X'X)^{-1} X'Y$$

CANPLS algorithm

$$\begin{aligned} &= \\ &= \sum \lambda \\ \parallel &= \end{aligned}$$

$$\begin{aligned} &= - \\ &= \sum \lambda \\ \parallel &= \end{aligned}$$

$$\beta = -$$

Modified Continuum power regression

$$\sim \infty \quad \gamma \quad = \sum_{=} \lambda^\gamma$$

$$\sim \infty \quad - \quad \gamma \quad - \quad = \sum_{=} \lambda^\gamma \quad -$$

Modified Continuum power regression

$$\sim \infty - \gamma - = \sum_{=} \lambda^\gamma -$$

$$\gamma =$$

$$\gamma =$$

$$\gamma \rightarrow \infty$$

=

=

=

=

Régression PLS et très grande dimension

Bush et Nachbar (1993)

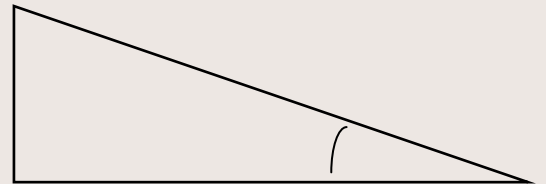
Sample-based PLS



$$\theta = \frac{+}{-}$$

$$= \theta = \frac{+}{-}$$

$$= \sqrt{\quad} \sqrt{\quad} = \quad ,$$



Sample-based PLS

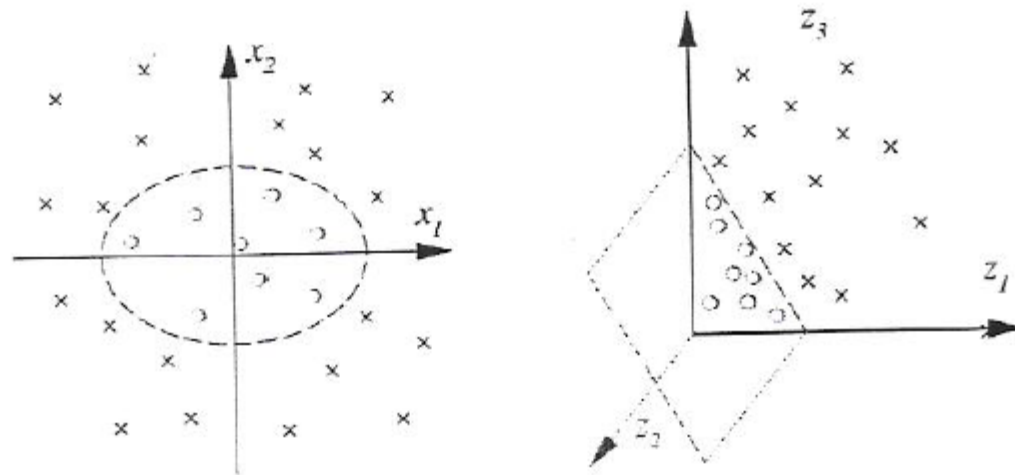
$$= \sum (\quad - \quad) \Rightarrow = - (\quad - \quad - \quad + \quad)$$
$$= \quad ,$$

→

Kernel methods

$$l = 2, p = 3$$

$$\phi : (x_1, x_2) \rightarrow (x_1^2, \sqrt{2}x_1x_2, x_2^2)$$



En 2D : une surface de décision non linéaire est nécessaire pour séparer les données

En 3D : un hyperplan est suffisant.

Kernel methods

$$\begin{aligned}
 &= \Phi = \sqrt{\quad} \\
 \langle \Phi \quad \Phi \rangle &= + \quad + \\
 &= + \quad = \langle \quad \rangle \\
 &= (\langle \quad \rangle +) \\
 &= \frac{\| \quad - \quad \|}{\sigma}
 \end{aligned}$$

Régression PLS et optimisation

$$\begin{aligned} &= \sum_{i=1}^n \lambda_i^{-1} \mathbf{v}_i \mathbf{v}_i^T \mathbf{X}^T \mathbf{Y} \\ &= \sum_{i=1}^n \lambda_i^{-1} \mathbf{v}_i \mathbf{v}_i^T \mathbf{X}^T \mathbf{Y} \end{aligned}$$

$$\begin{aligned} &= \sqrt{\Sigma} \mathbf{X}^T \mathbf{Y} \end{aligned} \rightarrow \begin{aligned} &= \sqrt{\Sigma} \mathbf{X}^T \mathbf{Y} \end{aligned}$$

Régression PLS généralisée

$$= \text{---} + \varepsilon \quad \longrightarrow \quad = \frac{\text{---}}{\text{---}}$$

$$= \quad + \quad + \quad + \quad - \quad - \quad + \quad -$$

$$= \quad + \quad + \quad + \quad - \quad - \quad + \quad \text{---} + \varepsilon \quad \longrightarrow \quad = \quad -$$

→

=

PLS-GLR algorithm



Available online at www.sciencedirect.com
SCIENCE @ DIRECT®
Computational Statistics & Data Analysis 48 (2005) 17–46

COMPUTATIONAL
STATISTICS
& DATA ANALYSIS

www.elsevier.com/locate/csda

PLS generalised linear regression

Philippe Bastien^a, Vincenzo Esposito Vinzi^{b,c,*}, Michel Tenenhaus^c

^a*L'Oréal Recherche, Aulnay, France*

^b*Department of Mathematics and Statistics University "Federico II", Naples, Italy*

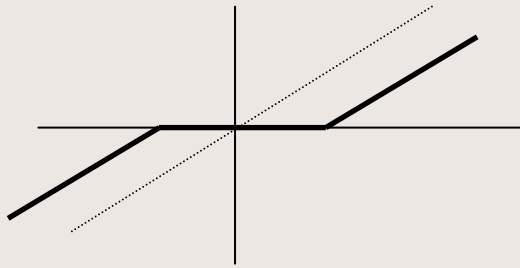
^c*HEC, School of Management Jussieu-Jesus, France*

Accepted 6 February 2004

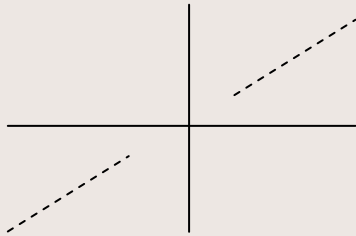
PLS-GLR and sparsity

Soft/Hard Thresholding

Soft thresholding



Hard thresholding



Régression logistique PLS

Table 8
Bordeaux wine data

OBS	Year	Temperature	Sunshine	Heat	Rain	Quality
1	1924	3064	1201	10	361	2
2	1925	3000	1053	11	338	3
3	1926	3155	1133	19	393	2
4	1927	3085	970	4	467	3
5	1928	3245	1258	36	294	1
6	1929	3267	1386	35	225	1
7	1930	3080	966	13	417	3
8	1931	2974	1189	12	488	3
9	1932	3038	1103	14	677	3
10	1933	3318	1310	29	427	2
11	1934	3317	1362	25	326	1
12	1935	3182	1171	28	326	3
13	1936	2998	1102	9	349	3
14	1937	3221	1424	21	382	1
15	1938	3019	1230	16	275	2
16	1939	3022	1285	9	303	2
17	1940	3094	1329	11	339	2
18	1941	3009	1210	15	536	3
19	1942	3277	1331	31	414	2
20	1943	3308	1366	24	282	1
21	1944	3212	1289	17	302	2
22	1945	3361	1444	25	253	1
23	1946	3061	1175	12	261	2
24	1947	3478	1317	42	259	1
25	1948	3126	1248	11	315	2
26	1949	3458	1508	43	286	1
27	1950	3252	1361	26	346	2
28	1951	3052	1186	14	443	3
29	1952	3270	1399	24	306	1
30	1953	3198	1259	20	367	1
31	1954	2904	1164	6	311	3
32	1955	3247	1277	19	375	1
33	1956	3083	1195	5	441	3
34	1957	3043	1208	14	371	3

The following variables (Table 8) were measured in 34 years (1924–1957):

Temperature:	Sum of average day temperatures (°C)	Temperature
Sunshine:	Duration of sunshine (h)	Sunshine
Heat:	Number of very warm days	Heat
Rain:	Rain height (mm)	Rain
Wine quality:	1 = good, 2 = average, 3 = poor	Wine quality

Ordinal logistic regression

VI

Ordinal logistic regression

Table 9
Logistic regression of quality on the standardised meteorological variables

Score test for the proportional odds assumption
Chi-square = 2.9159 with 4 DF ($p = 0.5720$)

Analysis of maximum likelihood estimates

Variable	DF	Parameter estimate	Standard error	Wald chi-square	Pr > chi-square
INTERCP1	1	-2.6638	0.9266	8.2641	0.0040
INTERCP2	1	2.2941	0.9782	5.4998	0.0190
TEMPERA	1	3.4268	1.8029	3.6125	0.0573
SUNSHINE	1	1.7462	1.0760	2.6335	0.1046
HEAT	1	-0.8891	1.1949	0.5536	0.4568
RAIN	1	-2.3668	1.1292	4.3931	0.0361

Table 10
Prediction quality of model (27) by using classical logistic regression

Quality observed counts	Predicted			Total
	1	2	3	
1	8	3	0	11
2	2	8	1	11
3	0	1	11	12
Total	10	12	12	34

Régression logistique PLS

=

=

= $\begin{matrix} \parallel & \parallel \\ \parallel & \parallel \end{matrix}$

= $_ ,$

$$= \frac{\begin{matrix} + & & + & & - \\ \sqrt{(\quad) + (\quad) + (\quad) + (-\quad)} \end{matrix}}{\quad}$$

=

+

+

-

=

Régression logistique PLS

Table 11
Results from logistic regression of quality on component t_1

Analysis of parameter	Maximum DF	Likelihood estimate	Estimates std.error	Chi-square	Pr > Chi sq
Intercept	1	-2.2650	0.8644	6.8662	0.0088
Intercept2	1	2.2991	0.8480	7.3497	0.0067
$t1$	1	2.6900	0.7155	14.1336	0.0002

Cross-table of observed and predicted quality

Quality observed Counts	Predicted			Total
	1	2	3	
1	9	2	0	11
2	2	8	1	11
3	0	1	11	12
Total	11	11	12	34

Régression logistique PLS

Détermination de t_2

$$\leq = \frac{\alpha + \beta + \beta}{+ \alpha + \beta + \beta}$$

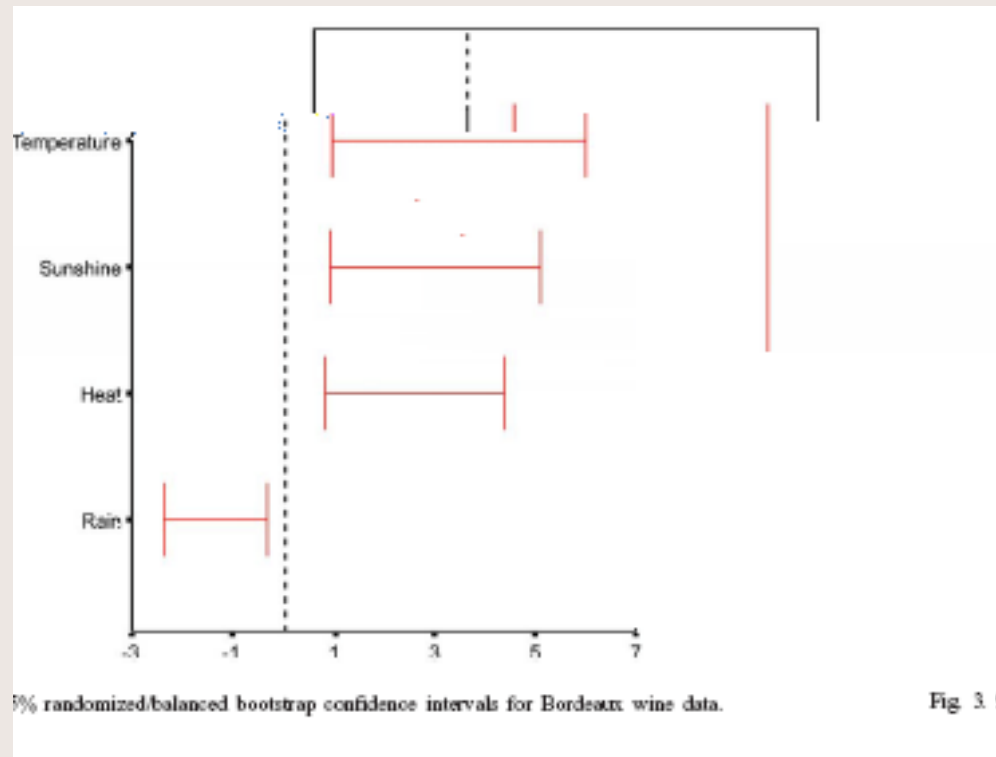
→

Expression du modèle final en fonction des descripteurs standardisés

$$= = \frac{- \quad + \quad \quad \quad + \quad \quad \quad + \quad \quad \quad -}{+ \quad - \quad + \quad \quad \quad + \quad \quad \quad + \quad \quad \quad -}$$

$$= \quad \quad \quad + \quad \quad \quad \quad \quad \quad + \quad \quad \quad \quad \quad -$$

Régression logistique PLS



The Cox proportional hazard model

$$\lambda = \lambda_0 e^{\beta'x}$$

$$\beta = \frac{\sum_{i \in D} x_i}{\sum_{i \in D} 1}$$

$$\beta$$

$$\leq$$

PLS-Cox regression

t

$$\begin{aligned} &= \\ &= \frac{\| \mathbf{X} \mathbf{t} \|}{\| \mathbf{X} \|} \\ &= \frac{\mathbf{X}^T \mathbf{X} \mathbf{t}}{\| \mathbf{X} \|^2} \end{aligned}$$

t

$$\begin{aligned} &= \\ &= \frac{\| \mathbf{X} \mathbf{t} \|}{\| \mathbf{X} \|} \\ &= \frac{\mathbf{X}^T \mathbf{X} \mathbf{t}}{\| \mathbf{X} \|^2} \end{aligned}$$

PLS-Cox regression

t_h

$$= \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}}$$

$$= \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}}$$

$$= \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}}$$

$$= \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2}}$$

||

||

Tibshirani (1997)

$$\beta = \beta \quad \sum |\beta| \leq$$

$$- \beta \quad - \beta \quad \sum |\beta| \leq$$

$$= \eta + \mu \eta \quad \beta \mu = \frac{\partial}{\partial \eta} \quad \frac{\partial}{\partial \eta \eta'}$$



Efron et al. (2004)

Gui and Li (2005)

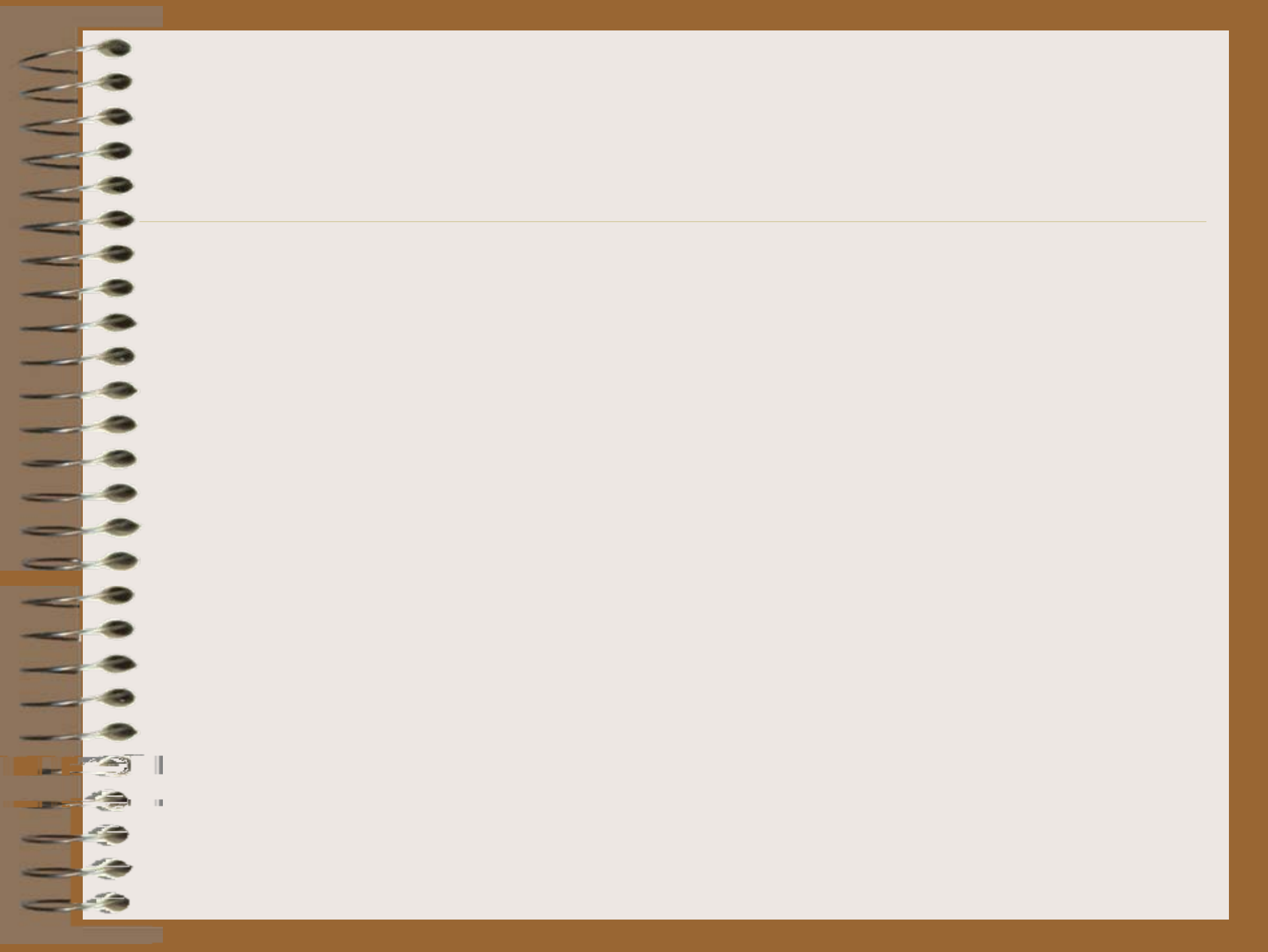
||

Segal (2006)

$$- \beta \quad - \beta \approx$$

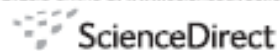
$$= \delta - \quad = \delta - \quad \beta'$$

$$= \quad - \quad - \delta \quad \frac{\delta -}{\delta}$$





Available online at www.sciencedirect.com



Chemometrics and Intelligent Laboratory Systems 91 (2006) 78–86

Chemometrics and
intelligent
laboratory systems

www.elsevier.com/locate/chemolab

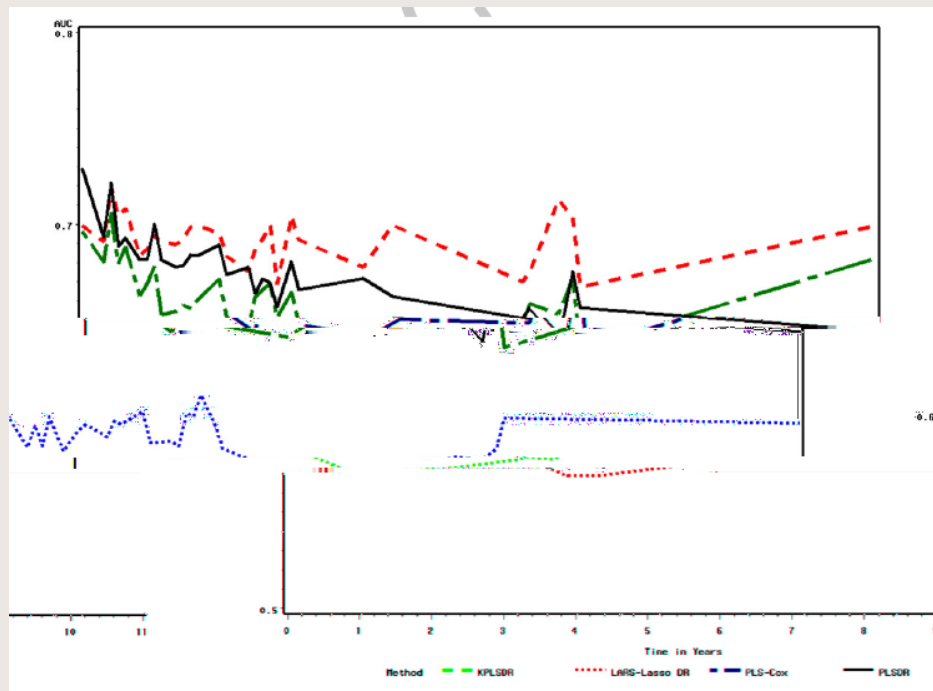
Deviance residuals based PLS regression for censored data in high dimensional setting

Philippe Bastien*

L'Oréal Recherche — Aulnay, 90 rue du Général Rogier, 92563 Clichy Cedex, France

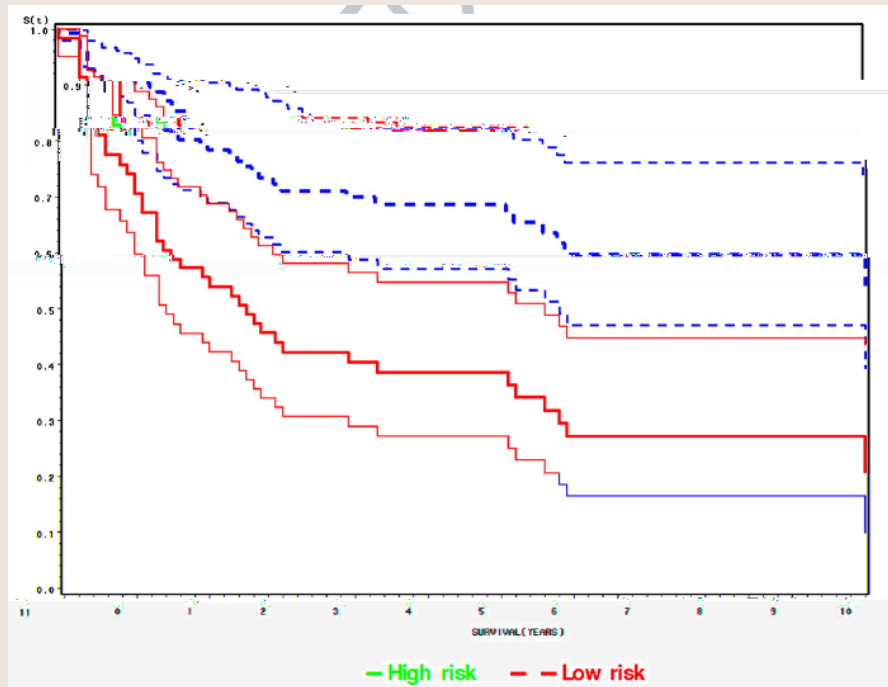
Received 29 December 2006; received in revised form 20 September 2007; accepted 20 September 2007
Available online 5 October 2007

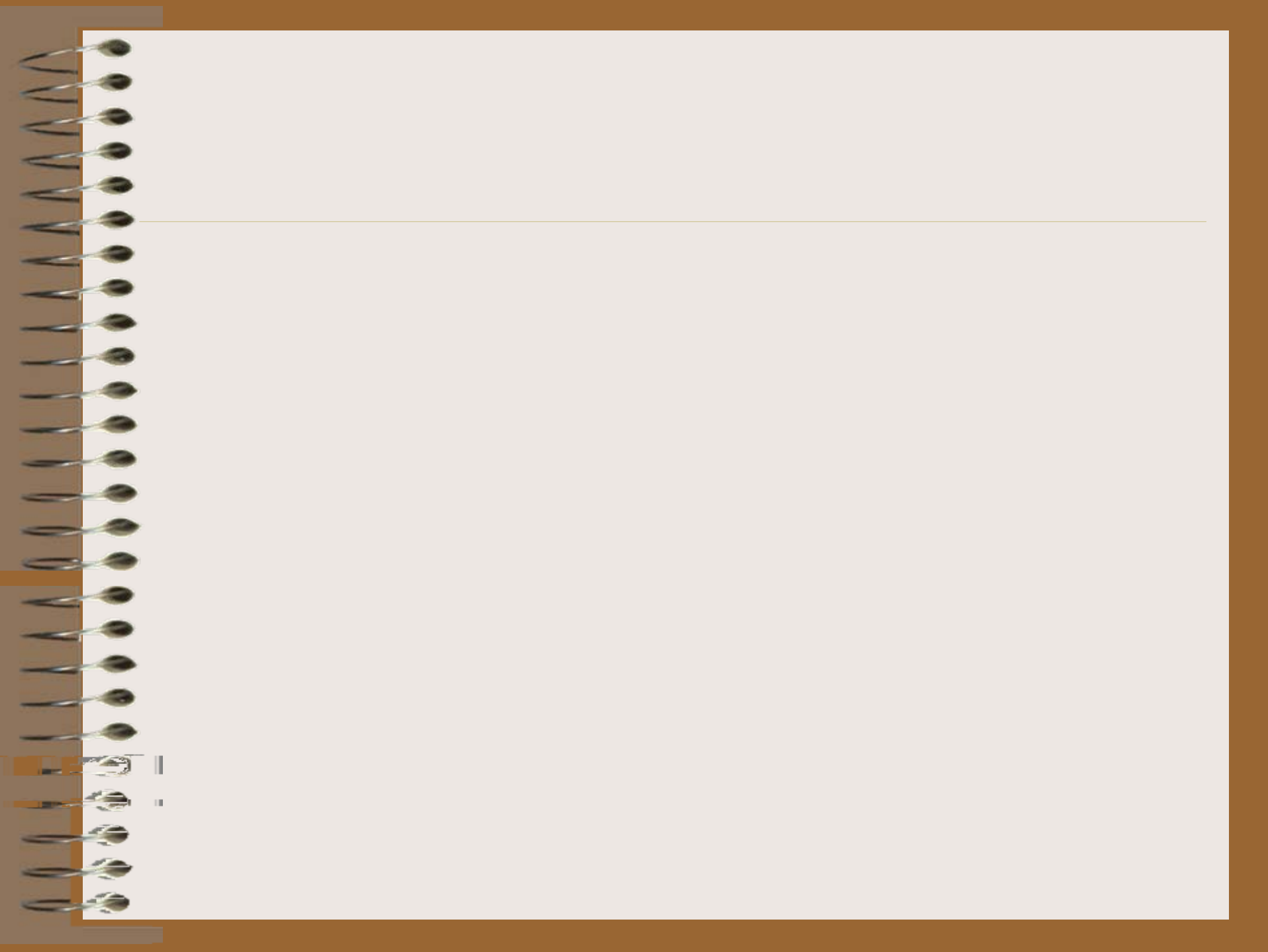
Application en transcriptomique



$\beta > \Rightarrow$

$\beta < \Rightarrow$





Packages sous R

‘integrOmics’ : R

‘plsRglm’ :

‘plsRcox’ :

PLSDOF :