Envelopes: Methods for Efficient Estimation in Multivariate Statistics

Dennis Cook

School of Statistics
University of Minnesota

LINSTAT2014

Collaborating with
Bing Li, Francesca Chiaromonte, Zhihua Su, Inge Helland &
Xin Zhang

Multivariate linear regression

$$\mathbf{Y}_i = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \ldots, n$$

- $\mathbf{Y} \in \mathbb{R}^r$: multivariate response
- $\mathbf{X} \in \mathbb{R}^p$:
  - Response reduction: non-stochastic predictors centered at 0
  - Predictor reduction: stochastic
- $\boldsymbol{\varepsilon} \in \mathbb{R}^r$: normal errors, mean 0 and covariance $\boldsymbol{\Sigma} > 0$
- $\boldsymbol{\alpha} \in \mathbb{R}^r$: unknown intercept
- $\boldsymbol{\beta} \in \mathbb{R}^{r \times p}$: unknown coefficients
- Goal: estimate $\boldsymbol{\beta}$, prediction.

MLE $\mathbf{B}$ of $\boldsymbol{\beta}$ is obtained by doing r univariate linear regressions, one for each response.

Rationale for envelopes

Envelopes arise by parameterizing the MLM in terms of the smallest subspace $\mathcal{E} \subseteq \mathbb{R}^r$ so that ($\mathbf{P}_\mathcal{E}$ = projection onto $\mathcal{E}$, $\mathbf{Q}_\mathcal{E} = \mathbf{I} - \mathbf{P}_\mathcal{E}$)

$$\begin{aligned} \mathbf{Q}_\mathcal{E}\mathbf{Y} \mid \mathbf{X} &\sim \mathbf{Q}_\mathcal{E}\mathbf{Y} \\ \mathbf{P}_\mathcal{E}\mathbf{Y} &\perp\!\!\!\perp \mathbf{Q}_\mathcal{E}\mathbf{Y} \mid \mathbf{X} \end{aligned}$$

This implies that the impact of $\mathbf{X}$ on $\mathbf{Y}$ is concentrated in $\mathbf{P}_\mathcal{E}\mathbf{Y}$. We refer to $\mathbf{P}_\mathcal{E}\mathbf{Y}$ and $\mathbf{Q}_\mathcal{E}\mathbf{Y}$ informality as the material and immaterial parts of $\mathbf{Y}$.

The conditions $\mathbf{Q}_{\mathcal{E}}\mathbf{Y} \mid \mathbf{X} \sim \mathbf{Q}_{\mathcal{E}}\mathbf{Y}$ and $\mathbf{P}_{\mathcal{E}}\mathbf{Y} \perp\!\!\!\perp \mathbf{Q}_{\mathcal{E}}\mathbf{Y} \mid \mathbf{X}$ hold if and only if

$$
\begin{aligned}
\operatorname{span}(\boldsymbol{\beta}) &\subseteq \mathcal{E} \\
\boldsymbol{\Sigma} &= \mathbf{P}_{\mathcal{E}}\boldsymbol{\Sigma}\mathbf{P}_{\mathcal{E}} + \mathbf{Q}_{\mathcal{E}}\boldsymbol{\Sigma}\mathbf{Q}_{\mathcal{E}}.
\end{aligned}
$$

- $\mathcal{E}$ envelops $\mathcal{B} := \operatorname{span}(\boldsymbol{\beta})$.
- $\mathcal{E}$ is a reducing subspace of $\boldsymbol{\Sigma}$.
- Formally, the intersection of all subspaces $\mathcal{E}$ with these properties is called the $\boldsymbol{\Sigma}$-envelope of $\mathcal{B}$ and represented as $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$ with $u = \dim(\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B}))$.
- Let the columns of the semi-orthogonal matrices $\boldsymbol{\Gamma} \in \mathbb{R}^{r \times u}$ and $\boldsymbol{\Gamma}_0 \in \mathbb{R}^{r \times (r-u)}$ be bases for $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$ and $\mathcal{E}^{\perp}_{\boldsymbol{\Sigma}}(\mathcal{B})$.

  Then $\boldsymbol{\beta} = \boldsymbol{\Gamma}\boldsymbol{\eta}$. $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Omega}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}_0\boldsymbol{\Omega}_0\boldsymbol{\Gamma}_0^T$, where $\boldsymbol{\Omega} > 0$ and $\boldsymbol{\Omega}_0 > 0$.

The envelope model becomes

$$\mathbf{Y} = \boldsymbol{\alpha} + \boldsymbol{\Gamma}\boldsymbol{\eta}\mathbf{X} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Omega}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}_0\boldsymbol{\Omega}_0\boldsymbol{\Gamma}_0^T.$$

Estimation via maximum likelihood with u determined by AIC, BIC, likelihood ratio testing, cross validation or a holdout sample.

We are still interested in $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$, which depend on the envelope $\mathcal{E}_{\boldsymbol{\Sigma}}(\mathcal{B})$, but not on the particular basis $\boldsymbol{\Gamma}$ selected to represent it. $\boldsymbol{\eta}$ and the $\boldsymbol{\Omega}$'s are basis dependent.

Envelope estimator $\widehat{\boldsymbol{\beta}} = \mathbf{P}_{\widehat{\mathcal{E}}}\mathbf{B}$, where $\mathbf{B}$ is the OLS estimator of $\boldsymbol{\beta}$.

How does envelope estimation work?

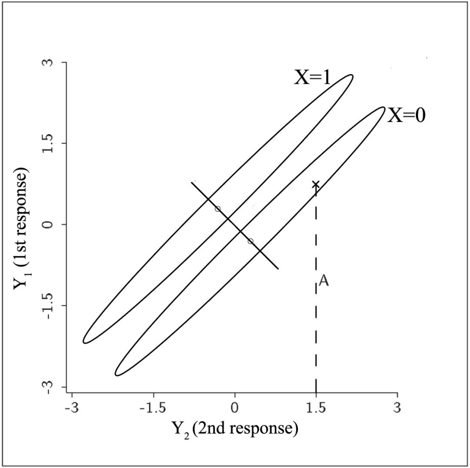
Multivariate regression with two responses, Y_1 and Y_2, and a single predictor, $X = 0$ or 1, to indicate two populations.

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} X + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

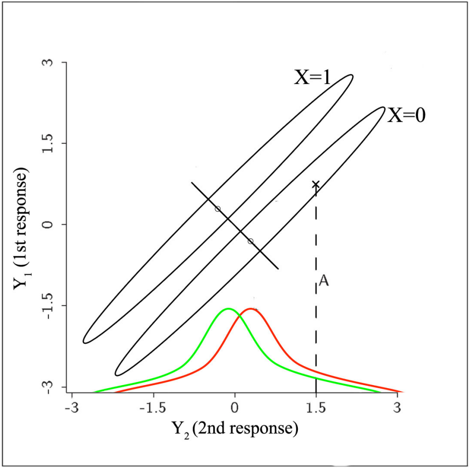
$\alpha_1 = E(Y_1|X = 0)$, $\beta_1 = E(Y_1|X = 1) - E(Y_1|X = 0)$,
$\alpha_2 = E(Y_2|X = 0)$, $\beta_2 = E(Y_2|X = 1) - E(Y_2|X = 0)$.

Standard estimators are obtained by substituting sample moments.

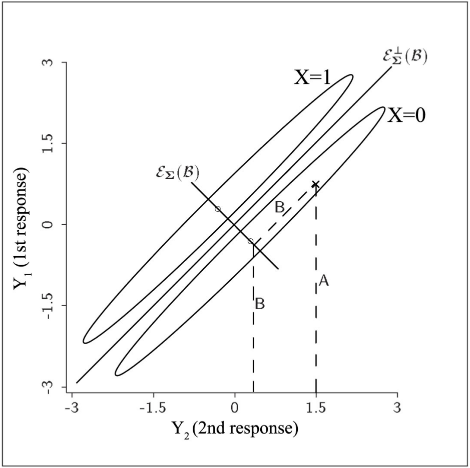
Schematic representation of standard analysis



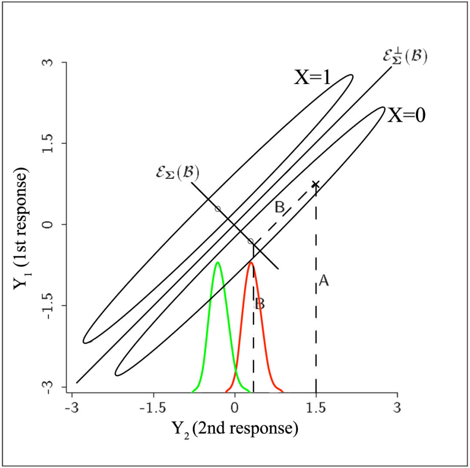
Schematic representation of standard analysis



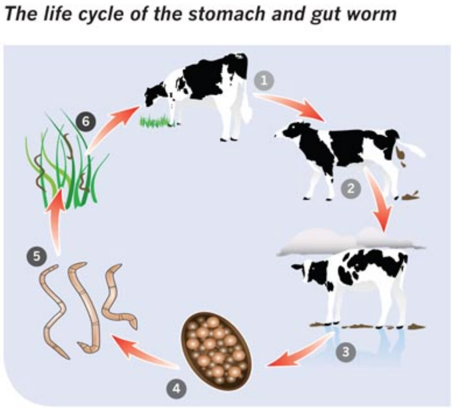
Working mechanism of envelope model



Working mechanism of envelope model

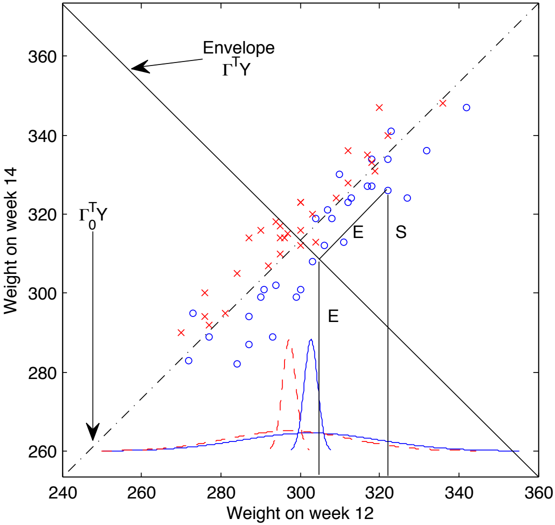


Cattle data



Experiment: Two treatments, each assigned randomly to 30 cows. Weight measured at weeks 2, 4, 6, . . . ,16, 18, 19.
Do the treatment have a differential effect; if so, about when it is apparent?

Cattle weight, week 12 vs week 14

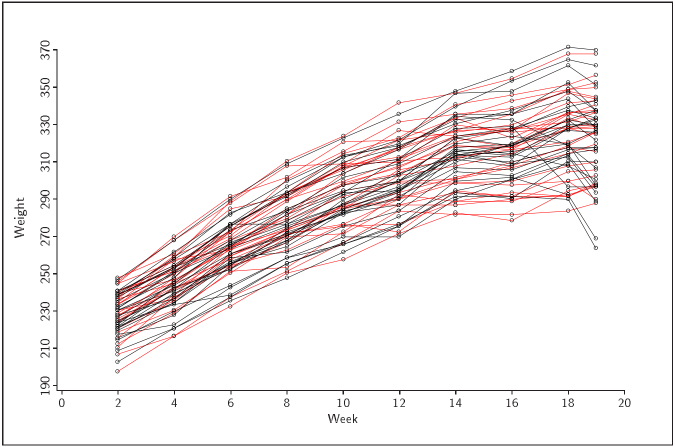


The OLS estimate is $\mathbf{B} = (5.5, -4.8)^T$ with bootstrap standard errors $(4.2, 4.4)^T$, while the envelope estimate is $\widehat{\beta} = (5.4, -5.1)^T$ with bootstrap standard errors $(1.12, 1.07)^T$.

About 1500 observations would be needed for an OLS analysis to yield the standard errors from an envelope analysis with 60 observations.

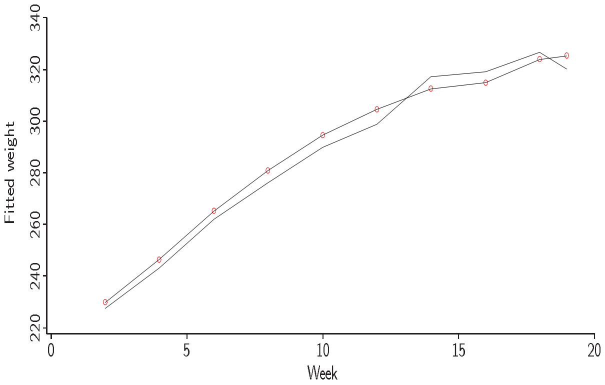
Next: Envelope analysis of the full data.

Profile plot of cattle data



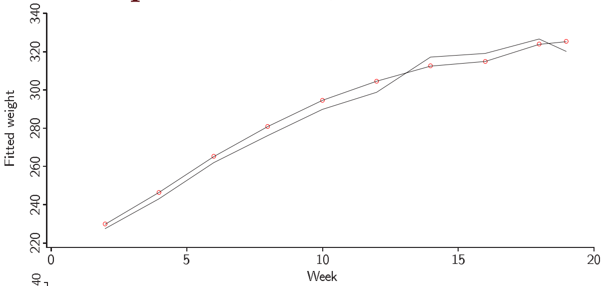
$$
\begin{aligned}
\mathbf{Y}_i &= \alpha + \beta X_i + \varepsilon_i, \quad X = 0, 1 \\
\mathbf{B} &= \text{OLS of } \beta = \bar{\mathbf{Y}}_{\text{trt1}} - \bar{\mathbf{Y}}_{\text{trt2}}
\end{aligned}
$$

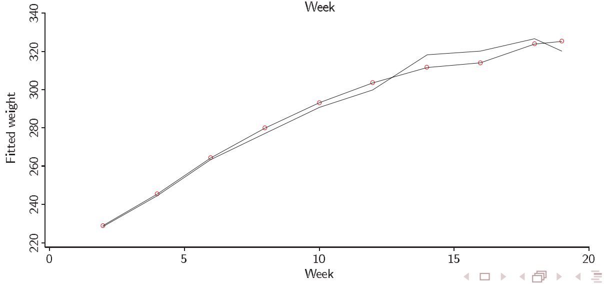
Mean profile plot of cattle data



$\max_i |B_i|/SE(B_i) \approx 1.3$. LRT stat. for $\boldsymbol{\beta} = 0$ is about 27 on 10 df.

Fitted profile plots, after inferring that $u = 5$. From envelope fit, $|\widehat{\beta}_i|/SE(\widehat{\beta}_i) > 4.1$ for $i \geqslant 10$.





Notes on Estimation

Maximum likelihood estimators

The estimated envelope $\widehat{\mathcal{E}}_{\mathbf{\Sigma}}(\mathcal{B})$ can be represented as

$$\widehat{\mathcal{E}}_{\mathbf{\Sigma}}(\mathcal{B}) = \arg\min_{\mathcal{S}}(\log|\mathbf{P}_{\mathcal{S}}\mathbf{S}_{\mathbf{Y}|\mathbf{X}}\mathbf{P}_{\mathcal{S}}|_0 + \log|\mathbf{Q}_{\mathcal{S}}\mathbf{S}_{\mathbf{Y}}\mathbf{Q}_{\mathcal{S}}|_0),$$

where $|\cdot|_0$ means the product of the non-zero eigenvalues, and $\mathcal{S}$ is a u-dim subspace of $\mathbb{R}^r$.

Estimators of other parameters:

- $\widehat{\boldsymbol{\beta}} = \mathbf{P}_{\widehat{\mathbf{\Gamma}}}\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}}$,
- $\hat{\boldsymbol{\eta}} = \widehat{\mathbf{\Gamma}}^T\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}}$.
- $\widehat{\mathbf{\Omega}} = \widehat{\mathbf{\Gamma}}^T\mathbf{S}_{\mathbf{Y}|\mathbf{X}}\widehat{\mathbf{\Gamma}}$, $\widehat{\mathbf{\Omega}}_0 = \widehat{\mathbf{\Gamma}}_0^T\mathbf{S}_{\mathbf{Y}}\widehat{\mathbf{\Gamma}}_0$.

Asymptotic variance of the MLE

$$\sqrt{n}[\mathrm{vec}(\widehat{\boldsymbol{\beta}}) - \mathrm{vec}(\boldsymbol{\beta})] \xrightarrow{\mathcal{D}} N_{rp}(0, \mathbf{V})$$

$$\begin{aligned} \mathbf{V} &= \mathrm{avar}\{\sqrt{n}\mathrm{vec}[\widehat{\boldsymbol{\beta}}]\} \\ &= \mathrm{avar}\{\sqrt{n}\mathrm{vec}[\widehat{\boldsymbol{\beta}}_{\boldsymbol{\Gamma}}]\} + \mathrm{avar}\{\sqrt{n}\mathrm{vec}[\mathbf{Q}_{\boldsymbol{\Gamma}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{\eta}}]\} \\ &\leqslant \mathrm{var}(\mathrm{vec}[\widehat{\boldsymbol{\beta}}_{\mathrm{OLS}}]) \end{aligned}$$

The efficiency gains can be massive, particularly when $\|\boldsymbol{\Omega}\| \ll \|\boldsymbol{\Omega}_0\|$. $\|\cdot\|$ = spectral norm

$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Omega}\boldsymbol{\Gamma}^T + \boldsymbol{\Gamma}_0\boldsymbol{\Omega}_0\boldsymbol{\Gamma}_0^T =$ material var. + immaterial var.

Illustrations

Air pollution data in Los Angeles

- 42 measurements at noon
- **Y**: measurements for CO, NO, NO2, O3 and HC.
- **X**: wind speed and solar radiation
- $\widehat{u} = 1$, $\|\widehat{\Omega}\| = 0.21$ and $\|\widehat{\Omega}_0\| = 36.3$.
- SE ratios for sm/em: $1.7 \sim 163$.

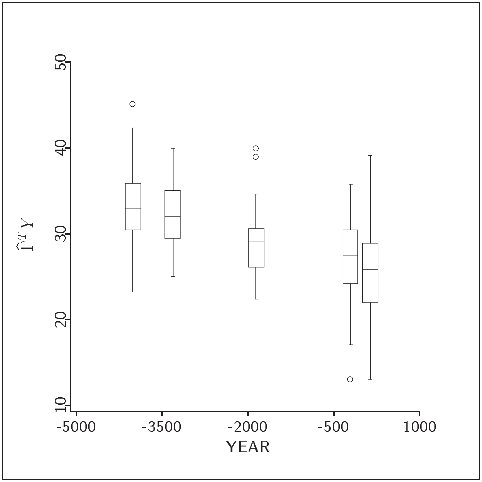
Individual SE ratios

4.3	5.7	CO
3.6	4.7	NO
51	68	NO2
123	163	O3
1.7	2.0	HC

Egyptian Skulls

- 4 measurements $\mathbf{Y}$ in cm on 30 male skulls in each of 5 epochs, 4000, 3300, 1850, 200 BC & 150 AD, included as indicators $\mathbf{X}$.
- $\mathbf{Y} = \alpha + \beta_{3300} X_1 + \beta_{1850} X_2 + \beta_{200} X_3 + \beta_{150} X_4 + \varepsilon$
- $\mathbf{Y} = \alpha + \Gamma\eta_{3300} X_1 + \Gamma\eta_{1850} X_2 + \Gamma\eta_{200} X_3 + \Gamma\eta_{150} X_4 + \varepsilon,$
- Since $\widehat{u} = 1$, Γ is 4×1 and we can plot $\widehat{\Gamma}^T \mathbf{Y}$ vs epoch.

Skull Boxplots vs Epoch



Reducing X and Partial least squares

PLS formulation

With $\mathbf{X}$ random we consider the same model

$$\mathbf{Y}_i = \boldsymbol{\alpha} + \boldsymbol{\beta}\mathbf{X}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \ldots, n,$$

but now the goal is to reduce $\mathbf{X}$. PLS operates by

1. Reducing $\mathbf{X} \rightarrow \widehat{\boldsymbol{\phi}}^T\mathbf{X}$ by using an iterative algorithm
2. Fitting $\mathbf{Y} = \boldsymbol{\alpha} + \boldsymbol{\eta}^T\{\widehat{\boldsymbol{\phi}}^T\mathbf{X}\} + \boldsymbol{\varepsilon}$ using OLS
3. Estimating $\widehat{\boldsymbol{\beta}}^T_{\text{pls}} = \widehat{\boldsymbol{\phi}}\hat{\boldsymbol{\eta}} = \mathbf{P}_{\widehat{\boldsymbol{\phi}}(\mathbf{S}_{\mathbf{X}})}\mathbf{B}^T$

SIMPLS algorithm for $\widehat{\Phi}$ (de Jong, 1993)

Set $\mathbf{w}_0 = 0$ and let $\widehat{\boldsymbol{\phi}}_k = (\mathbf{w}_0, \ldots, \mathbf{w}_k) \in \mathbb{R}^{p \times k}$. Then given $\widehat{\boldsymbol{\phi}}_k$, the next vector $\mathbf{w}_{k+1}$ is constructed as

$$\begin{aligned}
\mathcal{S}_k &= \operatorname{span}(\mathbf{S}_{\mathbf{X}} \widehat{\boldsymbol{\Phi}}_k) \\
\mathbf{w}_{k+1} &= \ell_{\max}(\mathbf{Q}_{\mathcal{S}_k} \mathbf{S}_{\mathbf{XY}} \mathbf{S}_{\mathbf{XY}}^T \mathbf{Q}_{\mathcal{S}_k}) \\
\widehat{\boldsymbol{\Phi}}_{k+1} &= (\mathbf{w}_0, \ldots, \mathbf{w}_k, \mathbf{w}_{k+1})
\end{aligned}$$

for $k = 1, \ldots, m-1$. m, the number of components, is chosen by cross-validation or a hold-out sample. Then $\widehat{\boldsymbol{\Phi}} = \widehat{\boldsymbol{\Phi}}_m$.

Envelope connection: With known m, $\operatorname{span}(\widehat{\boldsymbol{\phi}}_m)$ is a $\sqrt{n}$-consistent estimator of the $\boldsymbol{\Sigma}_{\mathbf{X}}$-envelope of $\operatorname{span}(\boldsymbol{\beta}^T)$, $\mathcal{E}_{\boldsymbol{\Sigma}_{\mathbf{X}}}(\mathcal{B}')$, where $\mathcal{B}' = \operatorname{span}(\boldsymbol{\beta}^T)$ and $m = \dim(\mathcal{E}_{\boldsymbol{\Sigma}_{\mathbf{X}}}(\mathcal{B}'))$.

Alternatively, we can use an envelope estimator for the same tasks:

$$\begin{aligned}
\mathbf{Y} &= \boldsymbol{\alpha} + \boldsymbol{\eta}^T\{\boldsymbol{\phi}^T\mathbf{X}\} + \boldsymbol{\varepsilon} \\
\boldsymbol{\Sigma}_{\mathbf{X}} &= \boldsymbol{\phi}\boldsymbol{\Delta}\boldsymbol{\phi}^T + \boldsymbol{\phi}_0\boldsymbol{\Delta}_0\boldsymbol{\phi}_0^T \\
\boldsymbol{\Sigma} &= \boldsymbol{\Sigma} \\
\widehat{\boldsymbol{\beta}} &= \mathbf{B}\mathbf{P}^T_{\widehat{\boldsymbol{\phi}}(\mathbf{S}_{\mathbf{X}})}
\end{aligned}$$

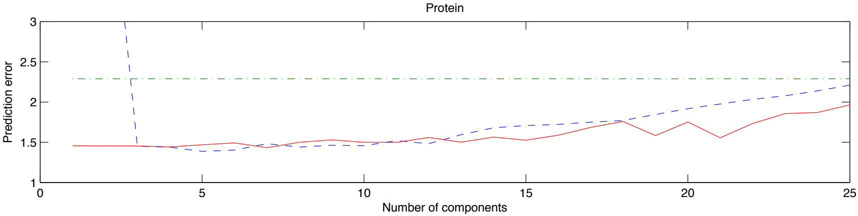
where

$$\widehat{\boldsymbol{\phi}} = \arg\min_{\mathcal{S}}\{\log|\mathbf{P}_{\mathcal{S}}\mathbf{S}_{\mathbf{X}|\mathbf{Y}}\mathbf{P}_{\mathcal{S}}|_0 + \log|\mathbf{Q}_{\mathcal{S}}\mathbf{S}_{\mathbf{X}}\mathbf{Q}_{\mathcal{S}}|_0\}$$

and $\mathcal{S}$ is an m-dim subspace of $\mathbb{R}^p$.

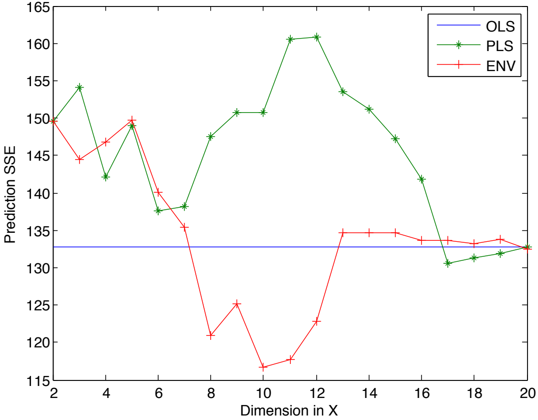
Beef protein

Predict protein content ($Y, r = 1$) of beef based on spectral measurements at $p = 50$ wave lengths, $n = 103$.



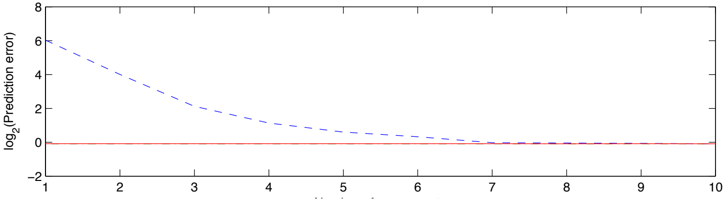
NIR analysis of biscuit dough

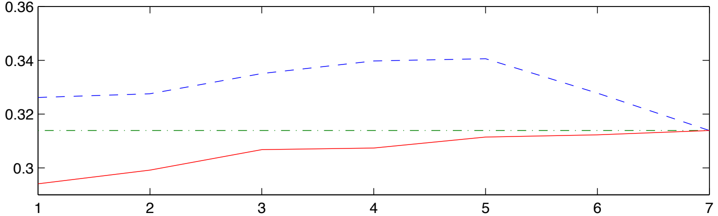
Predict fat, sucrose, flower and water content (Y, $r = 4$) of biscuit dough based on spectral measurements at $p = 20$ wave lengths, 39 training samples & 31 testing samples, created on different occasions. Comparison criterion is the SS prediction error on the testing samples.


165
160
155
150
145
140
135
130
125
120
115
Prediction SSE
2
4
6
8
10
12
14
16
18
20
Dimension in X
OLS
PLS
ENV

Simulations

Top: $r = 1, p = 10, u = 8$. $\mathbf{\Sigma_X} = 200\boldsymbol{\phi}\boldsymbol{\phi}^T + 50\boldsymbol{\phi}_0\boldsymbol{\phi}_0^T$
Bottom: $r = 1, p = 7, u = 2$. $\mathbf{\Sigma_X} = \boldsymbol{\phi}\boldsymbol{\Delta}\boldsymbol{\phi}^T + \boldsymbol{\phi}_0\boldsymbol{\Delta}_0\boldsymbol{\phi}_0^T$.
eigenvalues: 0.07 and 1.6 for $\boldsymbol{\Delta}$; between 3 and 584 for $\boldsymbol{\Delta}_0$.





Other envelope application in multivariate linear regression

- Partial response envelopes for part of $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$. (Su and Cook, *Biometrika*, 2011)

$$\begin{aligned}\mathbf{Y} &= \alpha + \boldsymbol{\beta}_1 \mathbf{X}_1 + \boldsymbol{\beta}_2 \mathbf{X}_2 + \varepsilon \\ &= \boldsymbol{\alpha} + \boldsymbol{\Gamma\eta} \mathbf{X}_1 + \boldsymbol{\beta}_2 \mathbf{X}_2 + \varepsilon \\ \boldsymbol{\Sigma} &= \boldsymbol{\Gamma\Omega\Gamma}^T + \boldsymbol{\Gamma}_0 \boldsymbol{\Omega}_0 \boldsymbol{\Gamma}_0^T\end{aligned}$$

- Simultaneous envelopes for reducing $\mathbf{X}$ and $\mathbf{Y}$ (Cook and Zhang, *Technometrics*, to appear)

$$\begin{aligned}\mathbf{Y} &= \alpha + \boldsymbol{\beta} \mathbf{X} + \varepsilon \\ &= \boldsymbol{\alpha} + \boldsymbol{\Gamma\eta\Phi}^T \mathbf{X} + \varepsilon \\ \boldsymbol{\Sigma} &= \boldsymbol{\Gamma\Omega\Gamma}^T + \boldsymbol{\Gamma}_0 \boldsymbol{\Omega}_0 \boldsymbol{\Gamma}_0^T \\ \boldsymbol{\Sigma}_{\mathbf{X}} &= \boldsymbol{\Phi\Delta\Phi}^T + \boldsymbol{\Phi}_0 \boldsymbol{\Delta}_0 \boldsymbol{\Phi}_0^T\end{aligned}$$

- Scaled predictor envelopes, when predictors are in different scales. (Su and Cook, submitted)

$$
\begin{aligned}
\mathbf{Y} &= \boldsymbol{\alpha} + \boldsymbol{\eta}^T \boldsymbol{\Phi}^T \boldsymbol{\Lambda}^{-1} \mathbf{X} + \boldsymbol{\varepsilon}, \\
\boldsymbol{\Sigma}_{\mathbf{X}} &= \boldsymbol{\Lambda} \boldsymbol{\Phi} \boldsymbol{\Delta} \boldsymbol{\Phi}^T \boldsymbol{\Lambda} + \boldsymbol{\Lambda} \boldsymbol{\Phi}_0 \boldsymbol{\Delta}_0 \boldsymbol{\Phi}_0^T \boldsymbol{\Lambda}, \\
\boldsymbol{\Lambda} &= \mathrm{diag}(1, \lambda_2, \ldots, \lambda_p)
\end{aligned}
$$

- Scaled response envelopes, when responses are in different scales. (Su and Cook, *Biometrika*, 2013)
- Inner envelopes, when envelopes don't offer improvement. (Su and Cook, *Biometrika*, 2012) – based on the largest reducing subspace of $\boldsymbol{\Sigma}$ that is contained within $\mathrm{span}(\boldsymbol{\beta})$.
- Heteroscedastic envelopes for comparing multivariate means in populations with different covariance matrices. (Su and Cook, *Statistica Sinica*, 2013).

Beyond linear models

Suppose we have an an asymptotically normal estimator $\widehat{\theta}$ of $\theta \in \mathbb{R}^p$, $\sqrt{n}(\widehat{\theta} - \theta) \to N(0, \mathbf{V}(\theta))$.

The estimator can often be improved by projecting it onto a root-n consistent estimator of the $\mathbf{V}(\theta)$-envelope of $\text{span}(\theta)$.

- Reproduces all of the known envelope methods, and applicable to GLMs.
- Links envelopes to a pre-specified estimator, MLE, robust estimator, OLS, ...
- $\mathbf{V}(\theta)$ can now depend on the parameter being estimated, plus perhaps nuisance parameters
- Don't need a likelihood to drive the process

Computing for linear model applications:

MatLab toolbox:
http://code.google.com/p/envlp/.

Thank you!

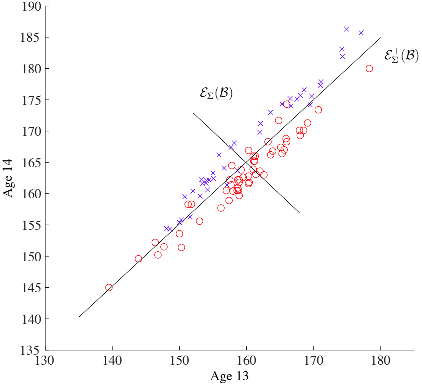
Table : Estimated coefficients from cattle data.

Week	$\mathbf{B}$	$\mathbf{B}/\text{se}(\mathbf{B})$	$\widehat{\boldsymbol{\beta}}$	$\widehat{\boldsymbol{\beta}}/\text{se}(\widehat{\boldsymbol{\beta}})$	$\text{se}(\mathbf{B})/\text{se}(\widehat{\boldsymbol{\beta}})$
2	2.43	0.83	-2.17	-1.67	2.25
4	3.33	1.05	-0.48	-0.65	4.27
6	3.13	0.89	0.88	1.23	4.89
8	4.73	1.22	2.38	2.82	4.56
10	4.73	1.14	2.89	4.14	5.94
12	5.50	1.30	5.40	5.30	4.15
14	-4.80	-1.11	-5.09	-5.55	4.69
16	-4.53	-0.97	-4.62	-5.36	5.40
18	-2.87	-0.54	-3.67	-4.06	5.86
19	5.00	0.86	4.21	4.92	6.78

We would need $n \sim 1500$ for OLS to match the envelope results.

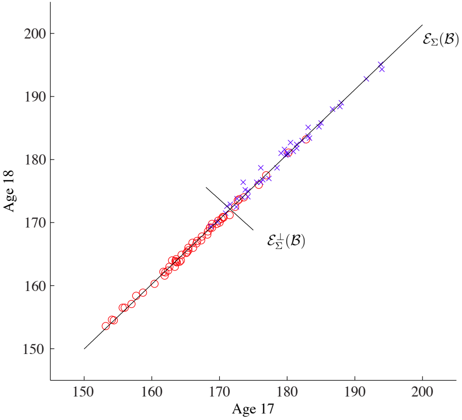
Heights of Boys and Girls

Heights of Boys and Girls Ages 13 and 14



- $\|\widehat{\Omega}\| = 1.57$ and $\|\widehat{\Omega}_0\| = 79.5$.
- SE ratios for sm/em: 8.49 and 8.61.

Heights of Boys and Girls Ages 17 and 18



- $\|\widehat{\Omega}\| = 118.7$ and $\|\widehat{\Omega}_0\| = 0.16$.
- SE ratios for sm/em: 1.01 and 0.99

Heights of Boys and Girls: Bootstrap SEs

Table : Bootstrap and estimated asymptotic standard errors of the two elements in $\widehat{\beta}$ under the standard model (SM) and envelope model (EM).

Response	SM	BSM	EM	BEM	SM/EM	BSM/BEM
Age 13	1.60	1.80	0.188	0.191	8.49	9.44
Age 14	1.61	1.81	0.187	0.190	8.61	9.64
Age 17	1.32	1.36	1.31	1.30	1.01	1.04
Age 18	1.33	1.37	1.34	1.37	0.99	1.01