Mathematics For Acousticians

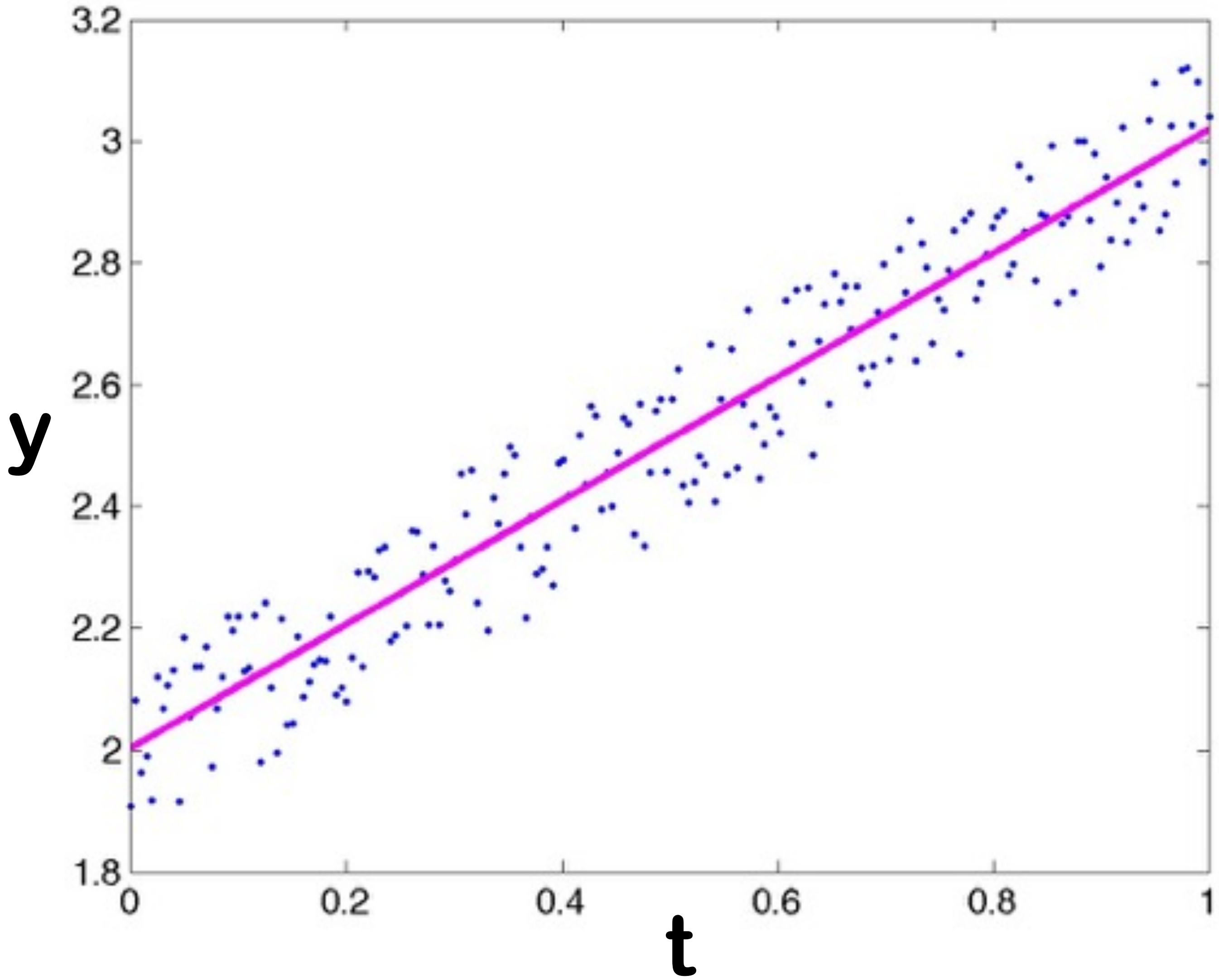
Least Mean Squares

Olivier Dazel
olivier.dazel@univ-lemans.fr
http://perso.univ-lemans.fr/~odazel/

Laboratoire d'Acoustique de l'Université du Maine
UMR CNRS 6613, France

MSc in Acoustics / IMDEA. Unit 167EN046

Example of a temporal sequence



Two vectors

$$\mathbf{t} = \begin{Bmatrix} t_1 \\ \vdots \\ t_n \end{Bmatrix} \quad \mathbf{y} = \begin{Bmatrix} y_1 \\ \vdots \\ y_n \end{Bmatrix}$$

$$f_{\alpha,\beta}(t) \approx \alpha t + \beta$$

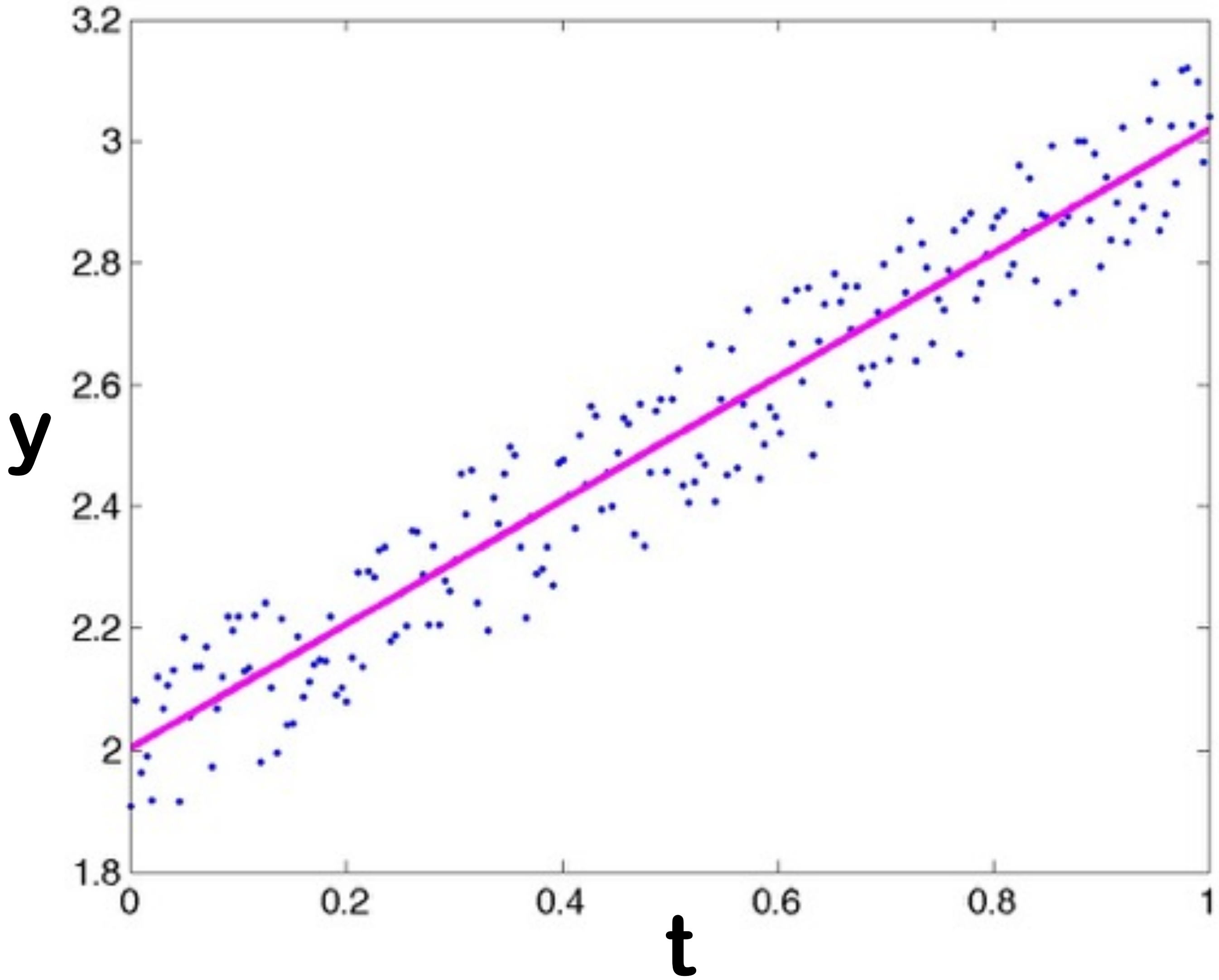
A criterion to minimize

$$\inf_{\boldsymbol{p}} \|f_{\mathbf{p}}(\mathbf{t}) - \mathbf{y}\|^2 = \inf_{\boldsymbol{p}} \sum_{k=1}^{n} \|f_{\mathbf{p}}(t_k) - y_k\|^2$$

Unknowns

Cost function

Example of a temporal sequence



Two vectors

$$\mathbf{t} = \left\{ \begin{array}{c} t_1 \\ \vdots \\ t_n \end{array} \right\} \quad \mathbf{y} = \left\{ \begin{array}{c} y_1 \\ \vdots \\ y_n \end{array} \right\}$$

$$f_{\alpha,\beta}(t) \approx \alpha t + \beta$$

$$\mathbf{p} = \left\{ \begin{array}{c} \alpha \\ \beta \end{array} \right\}$$

$$[\mathbf{M}] \in \mathcal{M}_{n,2}(\mathbb{R})$$

$$[\mathbf{M}] = [\mathbf{t}|\mathbf{1}]$$

$$\inf_{\boldsymbol{p}} \|[\mathbf{M}]\mathbf{p} - \mathbf{y}\|^2$$

Differential proof

$$\inf_{\mathbf{p}} \|[\mathbf{M}]\mathbf{p} - \mathbf{y}\|^2 = \inf_{\mathbf{p}} \underbrace{\left(\mathbf{p}^t[\mathbf{M}]^t[\mathbf{M}]\mathbf{p} - \mathbf{y}^t[\mathbf{M}]\mathbf{p} - \mathbf{p}^t[\mathbf{M}]^t\mathbf{y} + \mathbf{y}^t\mathbf{y}\right)}_{\mathcal{S}_{\mathbf{p}}}$$

Gradient of a symmetric quadratic form

$$\mathcal{S} = \mathbf{x}^t[\mathbf{A}]\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j \qquad [\mathbf{A}]^t = [\mathbf{A}] = [a_{ij}]$$

$$\frac{\partial \mathcal{S}}{\partial x_k} = 2a_{kk}x_k + \sum_{i\neq k} a_{ik}x_i + \sum_{j\neq k} a_{kj}x_j$$

$$= 2\sum_{i=1}^{n} a_{ik}x_i \qquad \boldsymbol{\nabla}\ \mathcal{S} = 2[\mathbf{A}]\mathbf{x}$$

Gradient of a linear form

$$\mathcal{S}' = \mathbf{v}^t\mathbf{x} = \sum_{i=1}^{n} v_ix_i \qquad \frac{\partial \mathcal{S}'}{\partial x_k} = v_k \qquad \boldsymbol{\nabla}\ \mathcal{S}' = \mathbf{v}$$

Differential proof

$$\inf_{\mathbf{p}} \|[\mathbf{M}]\mathbf{p}-\mathbf{y}\|^2 = \inf_{\mathbf{p}} \underbrace{\left(\mathbf{p}^t[\mathbf{M}]^t[\mathbf{M}]\mathbf{p}-\mathbf{y}^t[\mathbf{M}]\mathbf{p}-\mathbf{p}^t[\mathbf{M}]^t\mathbf{y}+\mathbf{y}^t\mathbf{y}\right)}_{\mathcal{S}_{\mathbf{p}}}$$

Expression of the gradient

$$\nabla S_{\mathbf{p}} = -2[\mathbf{M}]^t\mathbf{y} + 2[\mathbf{M}]^t[\mathbf{M}]\mathbf{p}$$

Linear system to solve

$$[\mathbf{M}]^t[\mathbf{M}]\mathbf{p} = [\mathbf{M}]^t\mathbf{y}$$

$[\mathbf{M}]^t[\mathbf{M}] \in \mathcal{M}_{(m\times m}(\mathbb{R})$

$[\mathbf{M}]^t\mathbf{y} \in \mathbb{R}^m$

Geometrical proof

$$\mathcal{I} = \{\mathbf{Z} \in \mathbb{R}^n \,|\, \exists \mathbf{x} \in \mathbb{R}^m, \quad \mathbf{Z} = [\mathbf{M}]\mathbf{x}\}$$

$$m = \dim(\mathcal{I}) \leqslant m$$

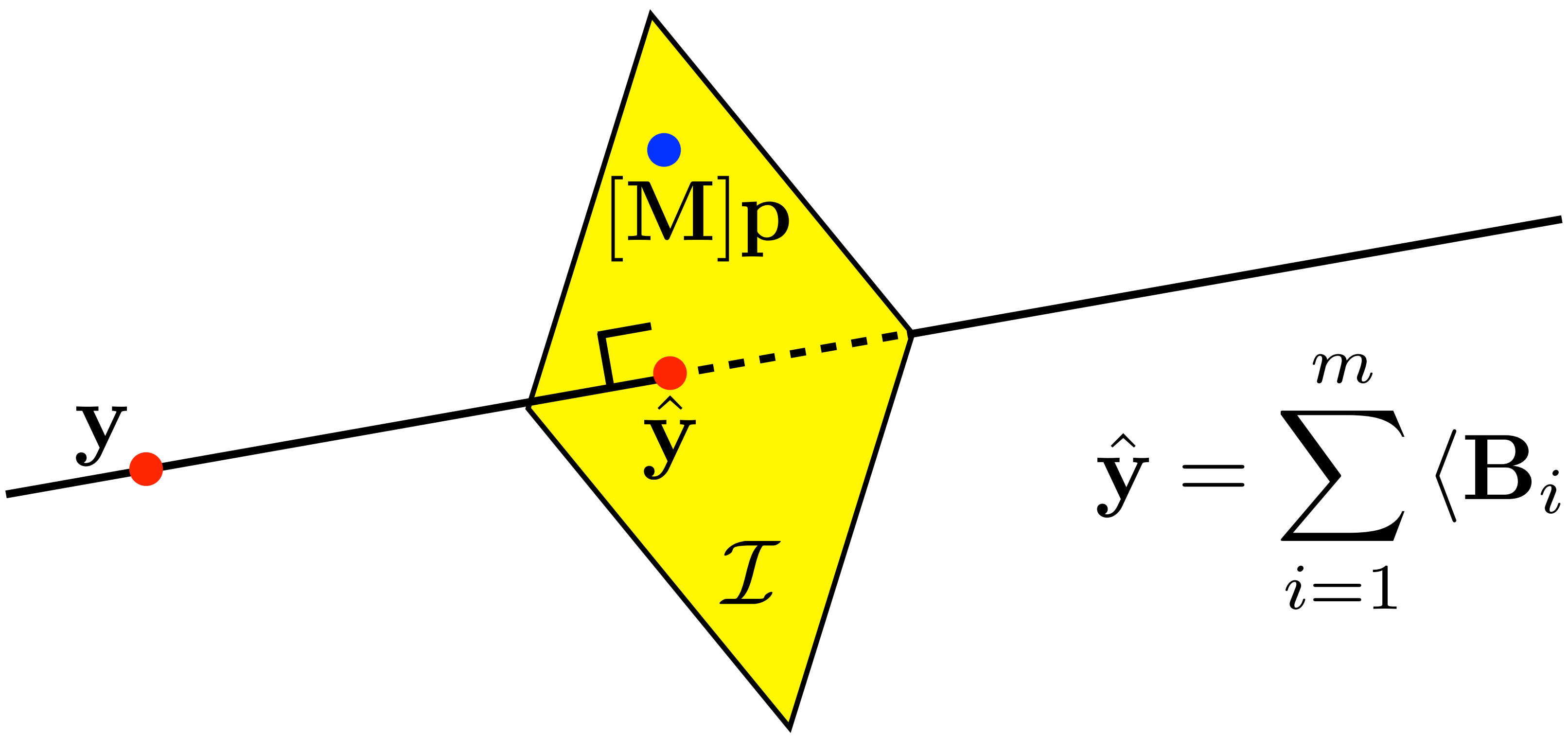
$$[\mathbf{B}] \in \mathcal{M}_{n,m}(\mathbb{R}) \quad [\mathbf{B}]^t[\mathbf{B}] = [\mathbf{I}_m]$$

Orthonormal basis of $\mathcal{I}$

$$[\mathbf{B}] = [\mathbf{M}][\mathbf{e}]$$

$$[\mathbf{e}] \in \mathcal{M}_{m,m}(\mathbb{R})$$

$$\det([\mathbf{e}]) \neq 0$$



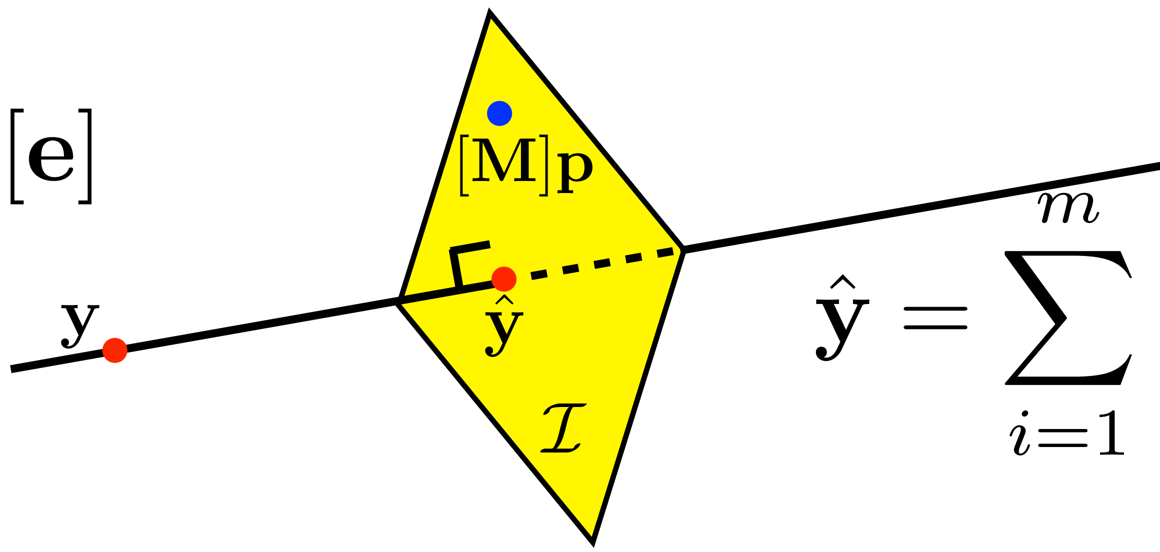
$$\hat{\mathbf{y}} = \sum_{i=1}^{m} \langle \mathbf{B}_i | \mathbf{y} \rangle \, \mathbf{B}_i = [\mathbf{B}][\mathbf{B}]^t \mathbf{y}$$

Pythagora

$$\|[\mathbf{M}]\mathbf{p} - \mathbf{y}\|^2 = \|[\mathbf{M}]\mathbf{p} - \hat{\mathbf{y}} - (\mathbf{y} - \hat{\mathbf{y}})\|^2$$
$$= \|[\mathbf{M}]\mathbf{p} - \hat{\mathbf{y}}\|^2 + \|\mathbf{y} - \hat{\mathbf{y}}\|^2$$

Geometrical proof

$$[\mathbf{B}] = [\mathbf{M}][\mathbf{e}]$$



$$\hat{\mathbf{y}} = \sum_{i=1}^{m} \langle \mathbf{B}_i | \mathbf{y} \rangle \, \mathbf{B}_i = [\mathbf{B}][\mathbf{B}]^t \mathbf{y}$$

$$\begin{aligned} \left\| [\mathbf{M}]\mathbf{p} - \mathbf{y} \right\|^2 &= \left\| [\mathbf{M}]\mathbf{p} - \hat{\mathbf{y}} - (\mathbf{y} - \hat{\mathbf{y}}) \right\|^2 \\ &= \left\| [\mathbf{M}]\mathbf{p} - \hat{\mathbf{y}} \right\|^2 + \left\| \mathbf{y} - \hat{\mathbf{y}} \right\|^2 \end{aligned}$$

We want to solve :

$$[\mathbf{M}]\mathbf{p} = \hat{\mathbf{y}}$$

$$[\mathbf{M}]\mathbf{p} = [\mathbf{B}][\mathbf{B}]^t \mathbf{y}$$

$$[\mathbf{B}]^t [\mathbf{M}]\mathbf{p} = [\mathbf{B}]^t \mathbf{y}$$

$$\cancel{[\mathbf{e}]^t} [\mathbf{M}]^t [\mathbf{M}]\mathbf{p} = \cancel{[\mathbf{e}]^t} [\mathbf{M}]^t \mathbf{y}$$

$$[\mathbf{B}]^t [\mathbf{B}] = [\mathbf{I}_m]$$

Validity of the approximation

Centering application:

$$c : \begin{cases} \mathbb{R}^n & \rightarrow & \mathbb{R}^n \\ \mathbf{x} & \mapsto & \mathbf{x} - \dfrac{<\mathbf{1},\mathbf{x}>}{n}\mathbf{1} \end{cases}$$

Average value

Normalization

$$\mathbf{t}' = \frac{c(\mathbf{t})}{\|c(\mathbf{t})\|}, \quad \mathbf{x}' = \frac{c(\mathbf{x})}{\|c(\mathbf{x})\|}$$

Correlation coefficient

$$R = <\mathbf{t}', \mathbf{x}'>$$

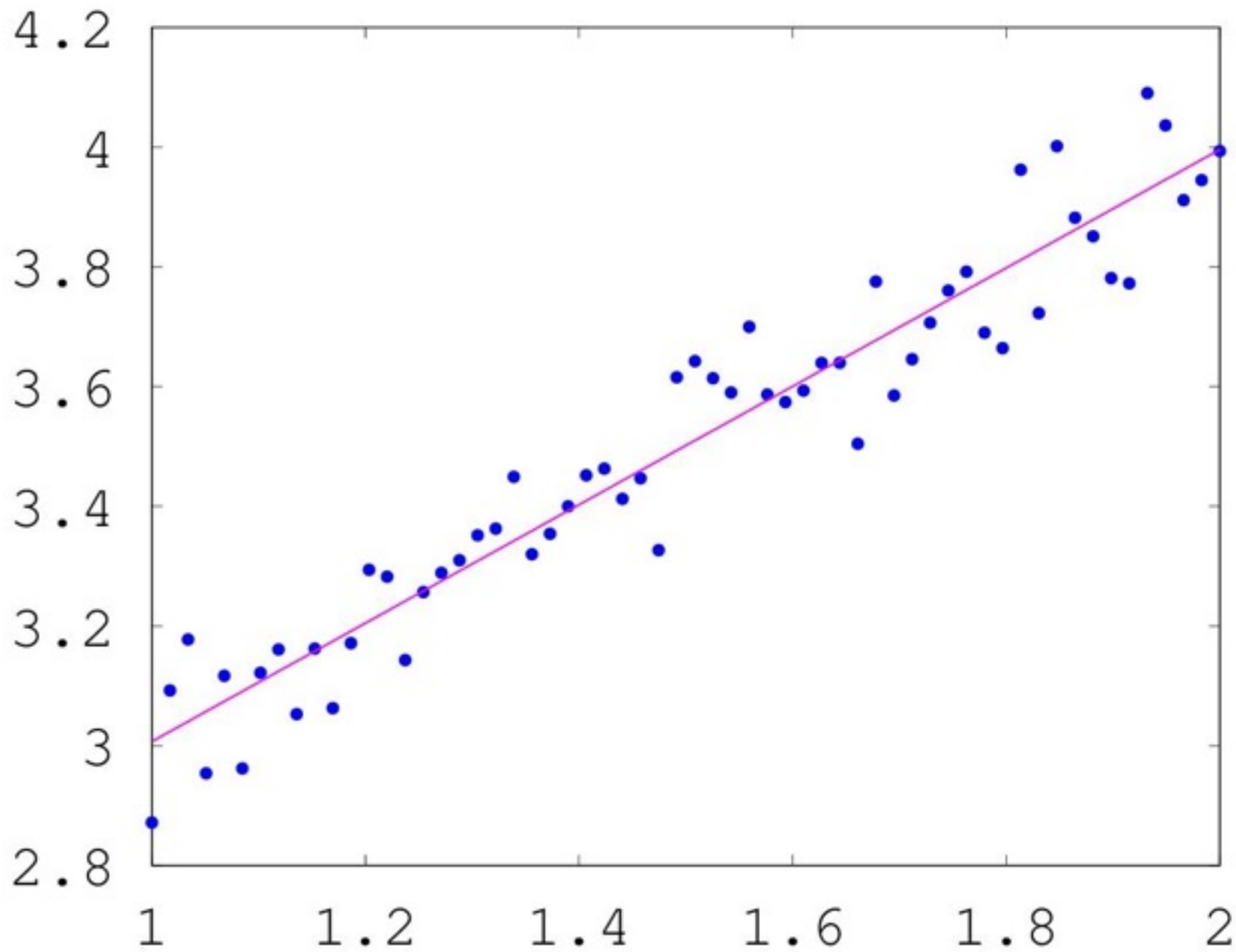
Example

$$y = \alpha t + \beta$$

$$\alpha = 1$$

$$\beta = 2$$

t: n values between 1 and 2



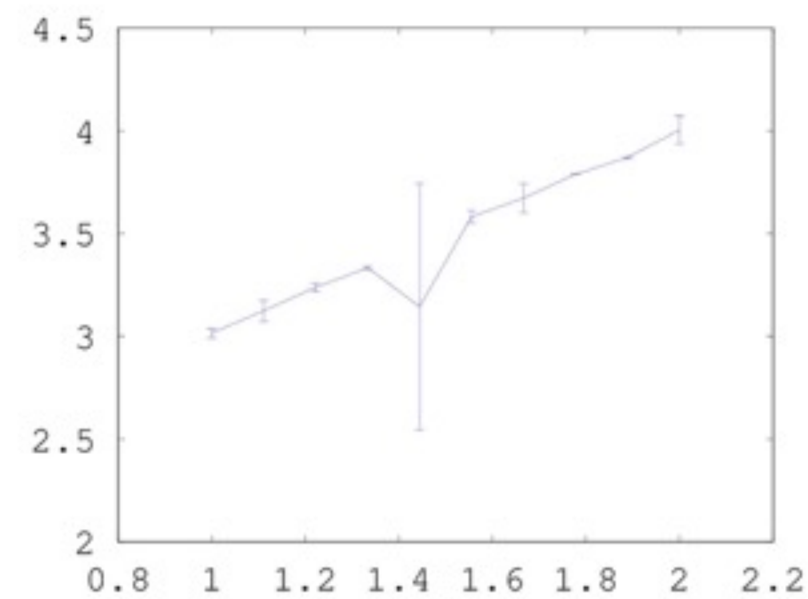
$$\alpha = 0.98127$$

$$\beta = 2.02420$$

$$R = 0.95630$$

Data with uncertainties

$$\mathbf{t} = \left\{ \begin{array}{c} t_1 \\ \vdots \\ t_n \end{array} \right\} \quad \mathbf{y} = \left\{ \begin{array}{c} y_1 \pm \sigma_1 \\ \vdots \\ y_i \pm \sigma_i \\ \vdots \\ y_n \pm \sigma_n \end{array} \right\}$$



Change of norm

$$\inf_{\mathbf{p}} \|[\mathbf{M}]\mathbf{p} - \mathbf{y}\|^2_{[\mathbf{w}]} = ([\mathbf{M}]\mathbf{p} - \mathbf{y})^t [\mathbf{w}] ([\mathbf{M}]\mathbf{p} - \mathbf{y})$$

Weight (ponderation) diagonal matrix

$$[\mathbf{w}] = \begin{bmatrix} \frac{1}{\sigma_1^2} & & & & \\ & \ddots & & & \\ & & \frac{1}{\sigma_i^2} & & \\ & & & \ddots & \\ & & & & \frac{1}{\sigma_n^2} \end{bmatrix}$$

Principle of diagonalization and SVD

$$\inf_{\mathbf{p}} \|[\mathbf{M}]\mathbf{p}-\mathbf{y}\|^2_{[\mathbf{w}]} = \inf_{\mathbf{p}} \left(\mathbf{p}^t[\mathbf{M}]^t[\mathbf{w}][\mathbf{M}]\mathbf{p} - \mathbf{y}^t[\mathbf{w}][\mathbf{M}]\mathbf{p} - \mathbf{p}^t[\mathbf{M}]^t[\mathbf{w}]\mathbf{y} + \mathbf{y}^t[\mathbf{w}]\mathbf{y}\right)$$

Decomposition of the weight matrix:

$$[\mathbf{w}] = \text{''}[\sqrt{\mathbf{w}}][\sqrt{\mathbf{w}}]\text{''}$$

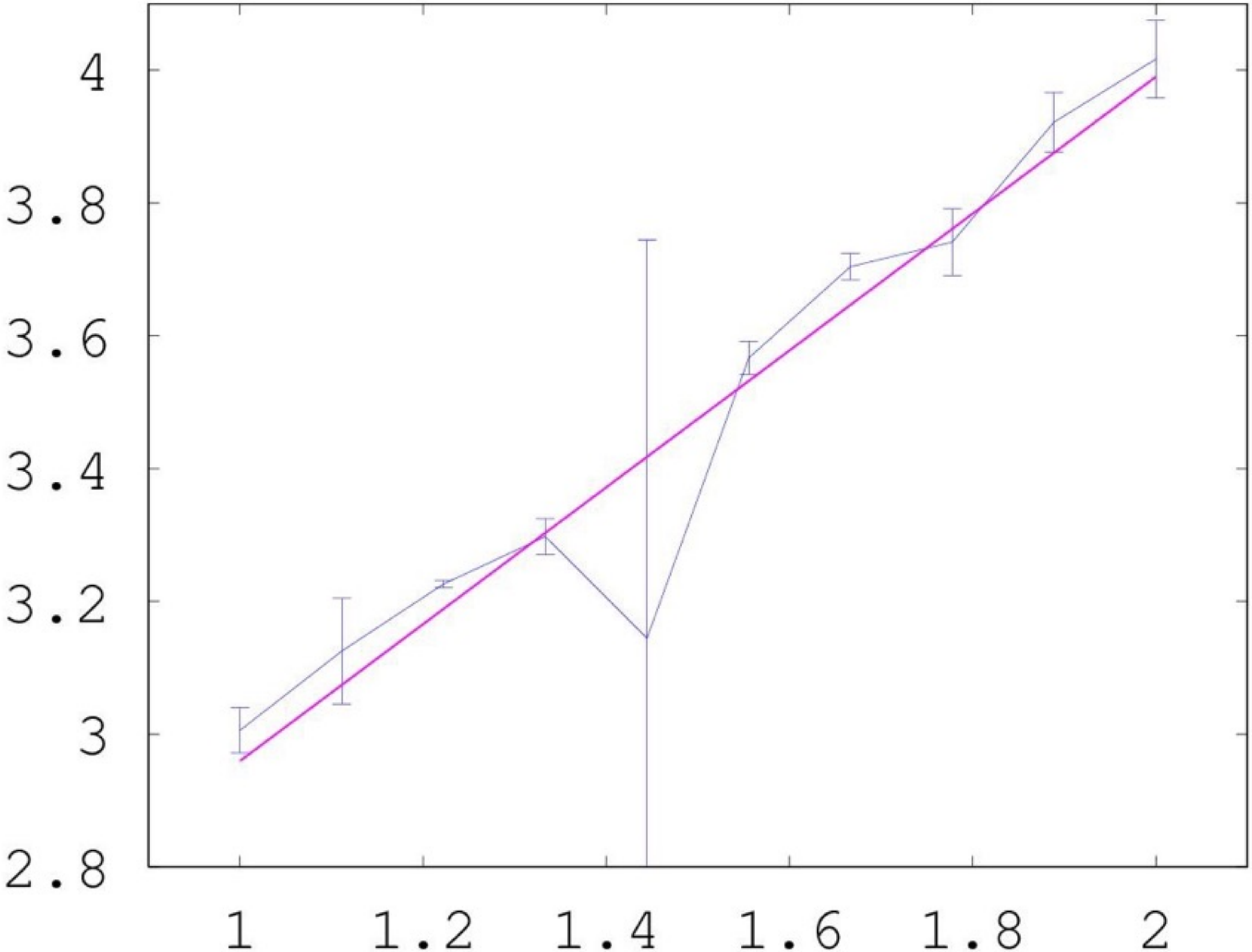
Ponderation of the problem matrices

$$[\mathbf{M}]' = [\sqrt{\mathbf{w}}][\mathbf{M}] \qquad \mathbf{y}' = [\sqrt{\mathbf{w}}]\mathbf{y}$$

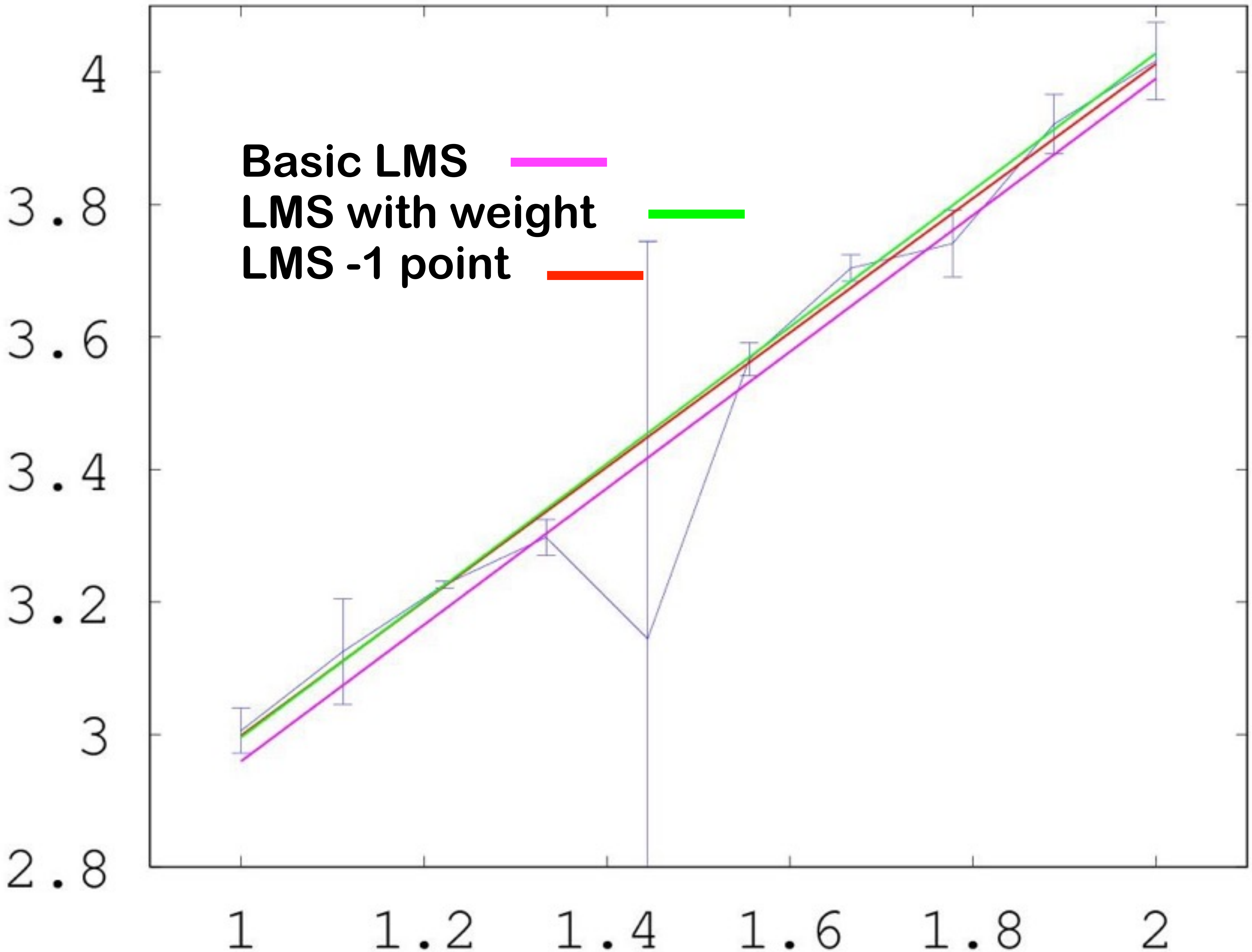
Back to the initial form

$$\begin{aligned}\inf_{\mathbf{p}} \|[\mathbf{M}]\mathbf{p}-\mathbf{y}\|^2_{[\mathbf{w}]} &= \inf_{\mathbf{p}} \left(\mathbf{p}^t[\mathbf{M}']^t[\mathbf{M}']\mathbf{p} - \mathbf{y}'^t[\mathbf{M}']\mathbf{p} - \mathbf{p}^t[\mathbf{M}']^t\mathbf{y}' + \mathbf{y}'^t\mathbf{y}'\right) \\ &= \inf_{\mathbf{p}} \left\|[\mathbf{M}']\mathbf{p}-\mathbf{y}'\right\|^2\end{aligned}$$

Example



Example



Basic LMS
LMS with weight
LMS -1 point
4
3.8
3.6
3.4
3.2
3
2.8
1
1.2
1.4
1.6
1.8
2