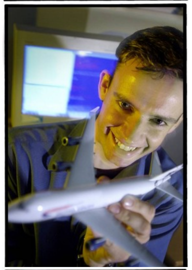





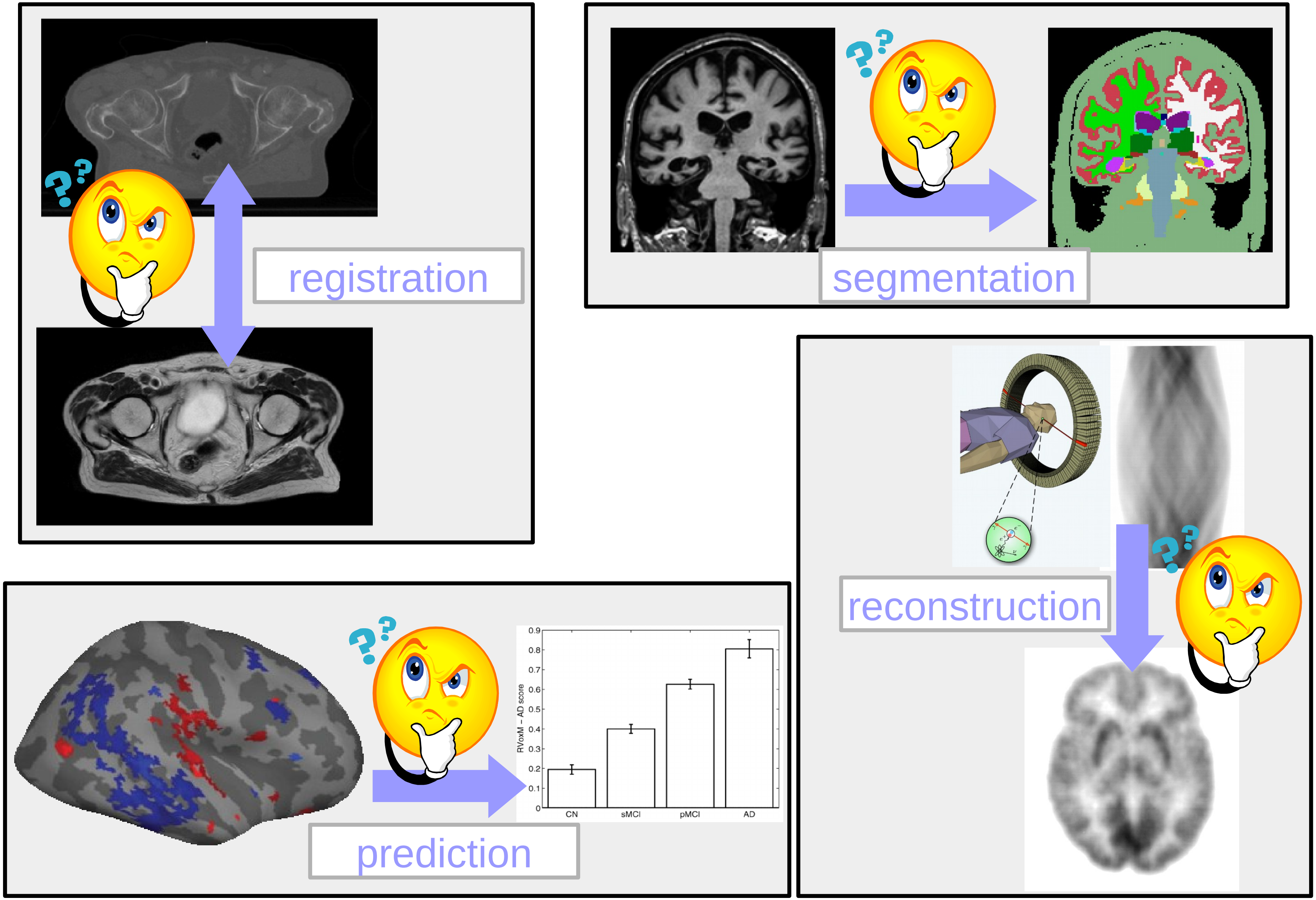

Medical Image Analysis

Course 22525

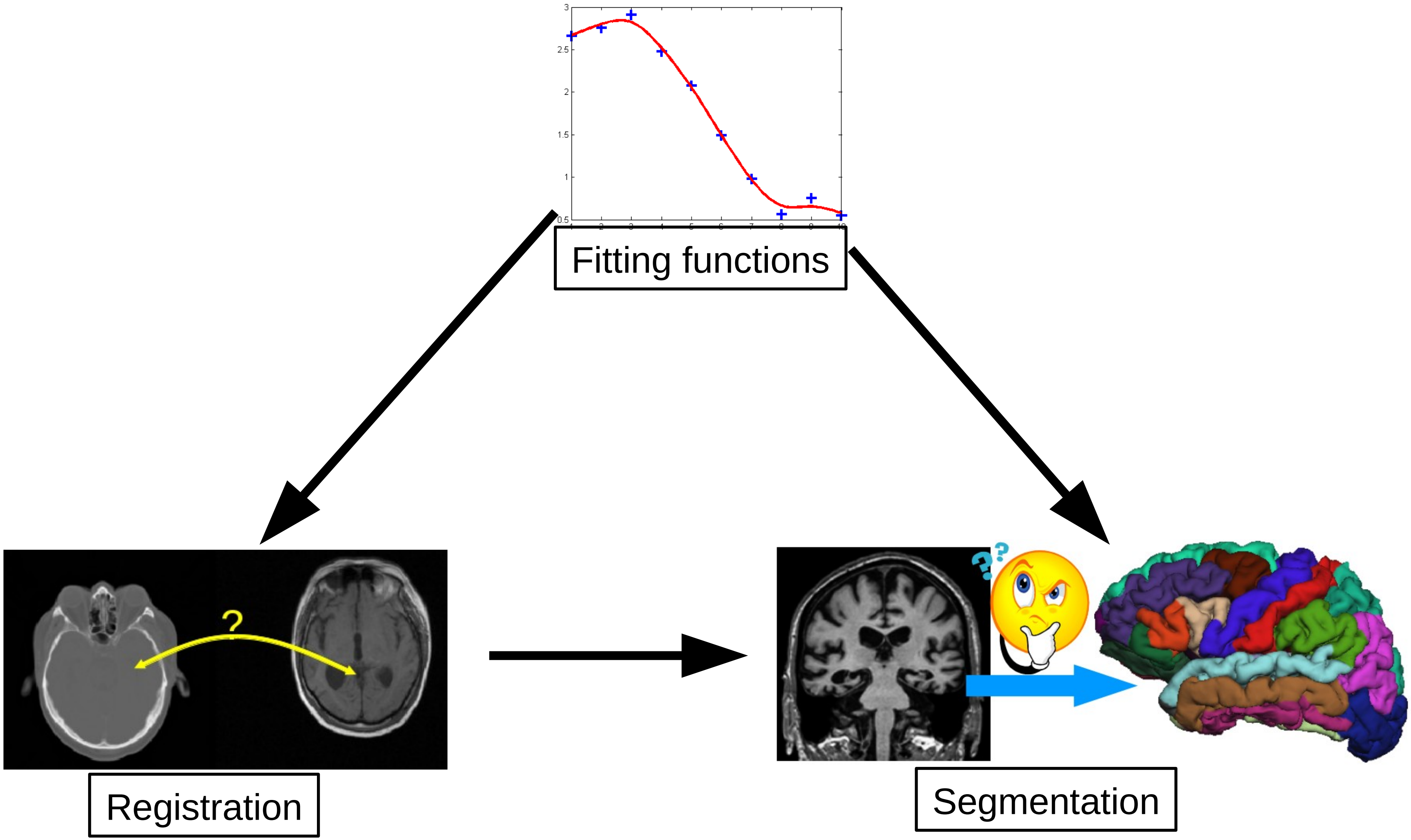
Koen Van Leemput

DTU HealthTech

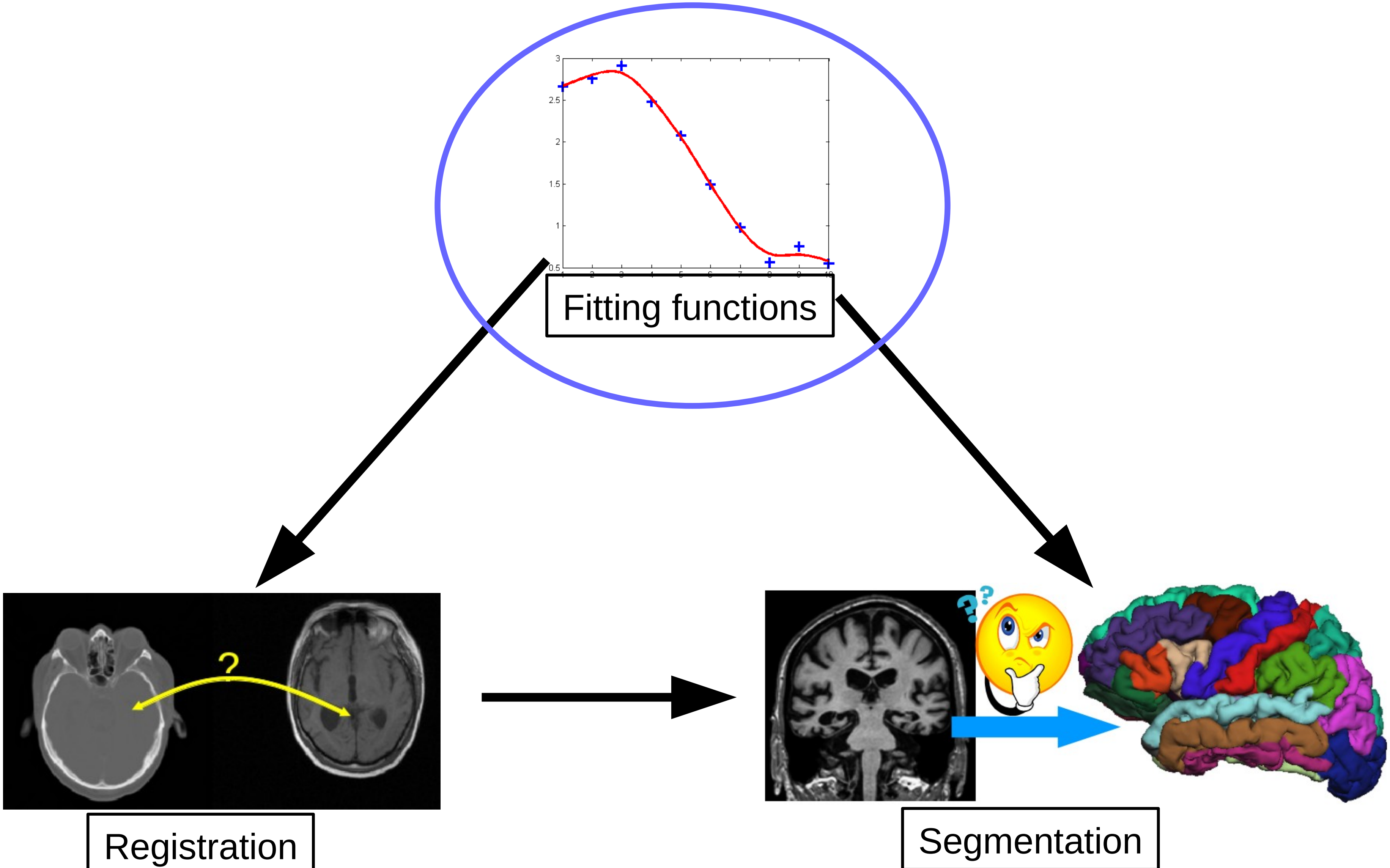
Technical University of Denmark


registration
segmentation
reconstruction
prediction
RVoxM – AD score
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
CN
sMCI
pMCI
AD

Course structure



Course structure



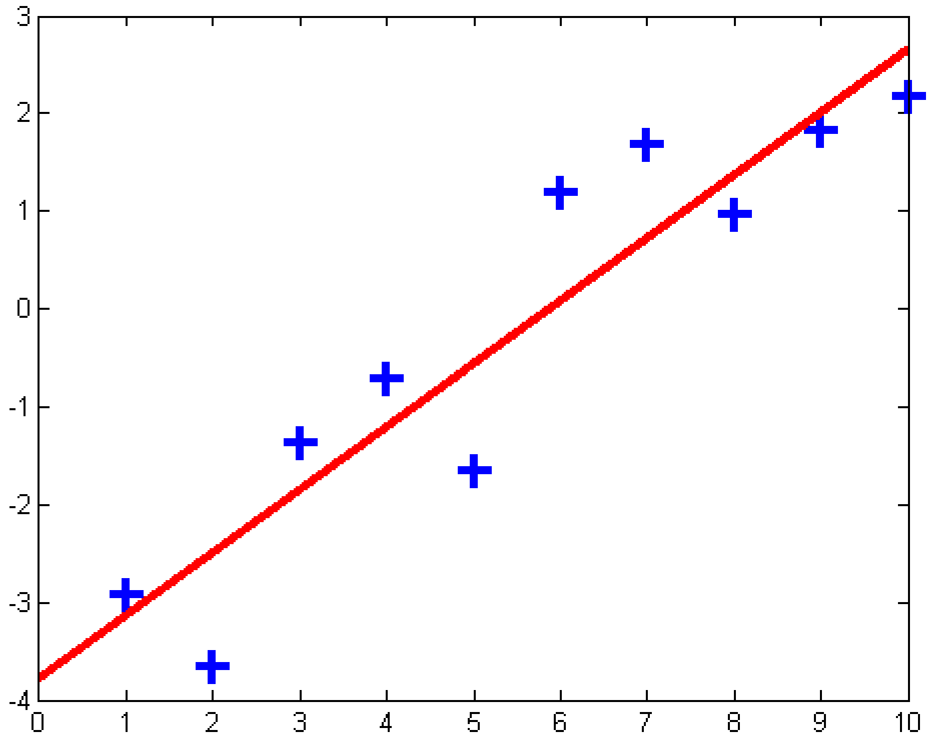
Linear regression

Input: $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^{\mathrm{T}} \in \mathbb{R}^p$

Observations: $\{\mathbf{x}_i, y_i\}_{i=1}^{N}$

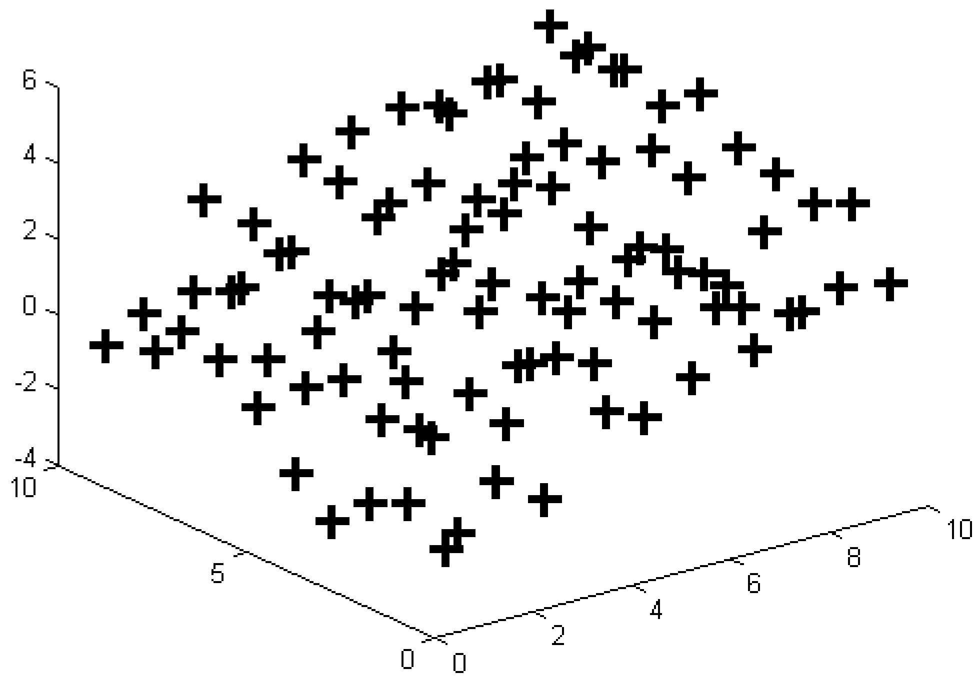
Linear regression model: $f(\mathbf{x}_i) = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j$

Example



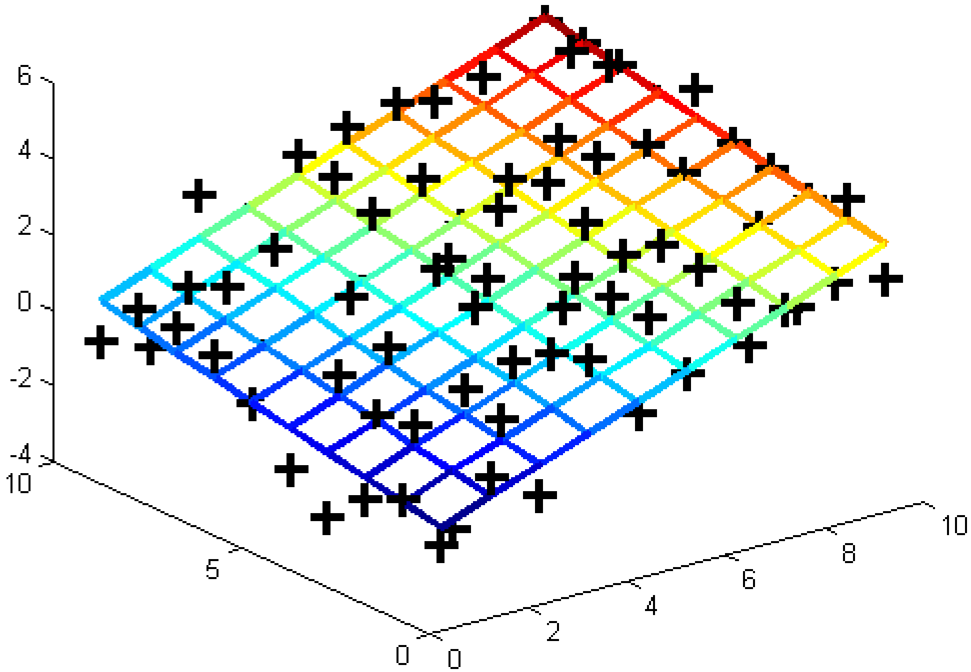
$$f(\mathbf{x}):\ \mathbf{x} \in \mathbb{R}$$

Example



$$f(\mathbf{x}): \mathbf{x} \in \mathbb{R}^2$$

Example



$$f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^2$$

Linear regression

Find parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^{\mathrm{T}}$ that minimizes

$$\mathrm{RSS}(\boldsymbol{\beta}) = \sum_{i=1}^{N} (y_i - f(\mathbf{x}_i))^2 = \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2$$

Matrix notation:

$$\mathrm{RSS} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & x_{N2} & \cdots & x_{Np} \end{pmatrix} \qquad \mathbf{y} = (y_1, y_2, \ldots, y_N)^{\mathrm{T}}$$

Linear regression

$$\mathrm{RSS} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

$$\frac{\partial \mathrm{RSS}}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Minimum found at: $\mathbf{X}^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = 0$

Linear regression

$$\mathbf{X}^{\mathrm{T}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = 0 \quad \Rightarrow \quad \hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$$

The predicted values at the training points are

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$$

$\mathbf{H} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$ is called the “hat” matrix

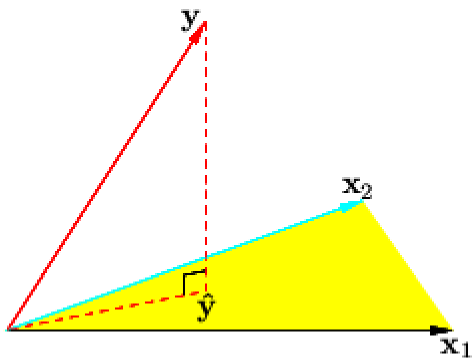


Figure 3.2: *The N-dimensional geometry of least squares regression with two predictors. The outcome vector $\mathbf{y}$ is orthogonally projected onto the hyperplane spanned by the input vectors $\mathbf{x}_1$ and $\mathbf{x}_2$. The projection $\hat{\mathbf{y}}$ represents the vector of the least squares predictions*

Residuals are orthogonal to column space of $\boldsymbol{X}$

$$\boldsymbol{X}^T(\boldsymbol{y} - \boldsymbol{X}\widehat{\beta}) = 0$$
$$\boldsymbol{X}^T(\boldsymbol{y} - \hat{\boldsymbol{y}}) = 0$$

$\boldsymbol{H}$ computes this projection.

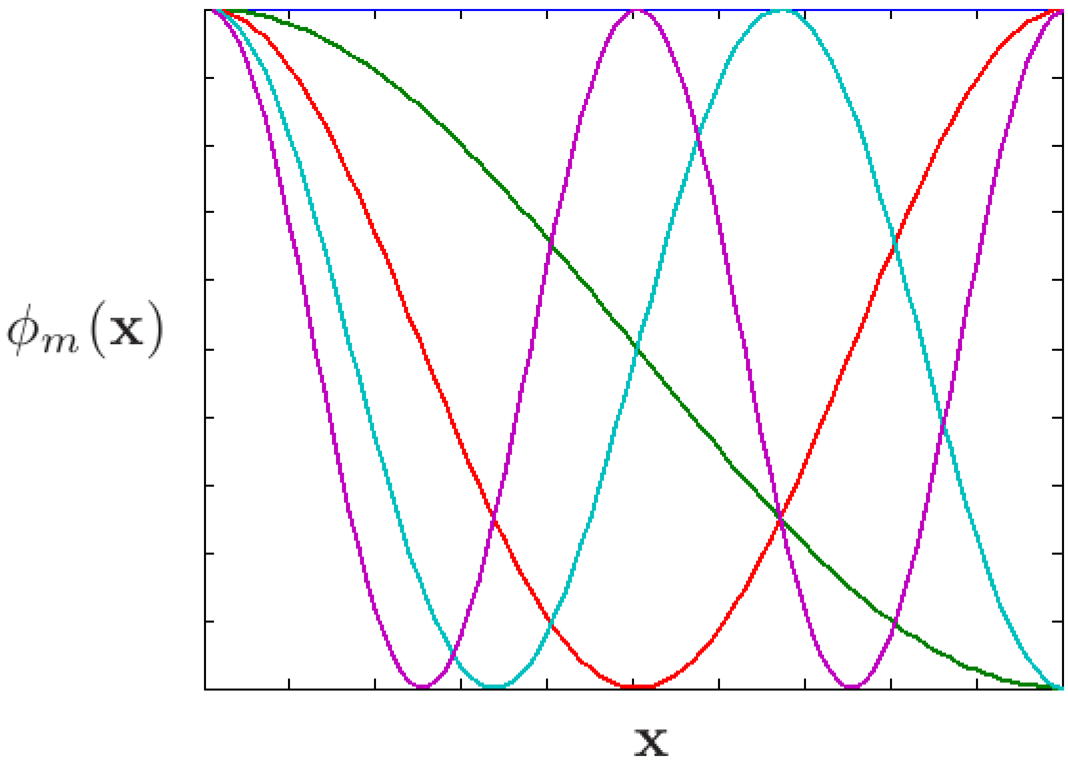
Linear basis function models

Models so far linear in both parameters β ***and*** input variables x_{ip}

More flexible models: $$f(\mathbf{x}) = \sum_{m=1}^{M} \beta_m \phi_m(\mathbf{x})$$

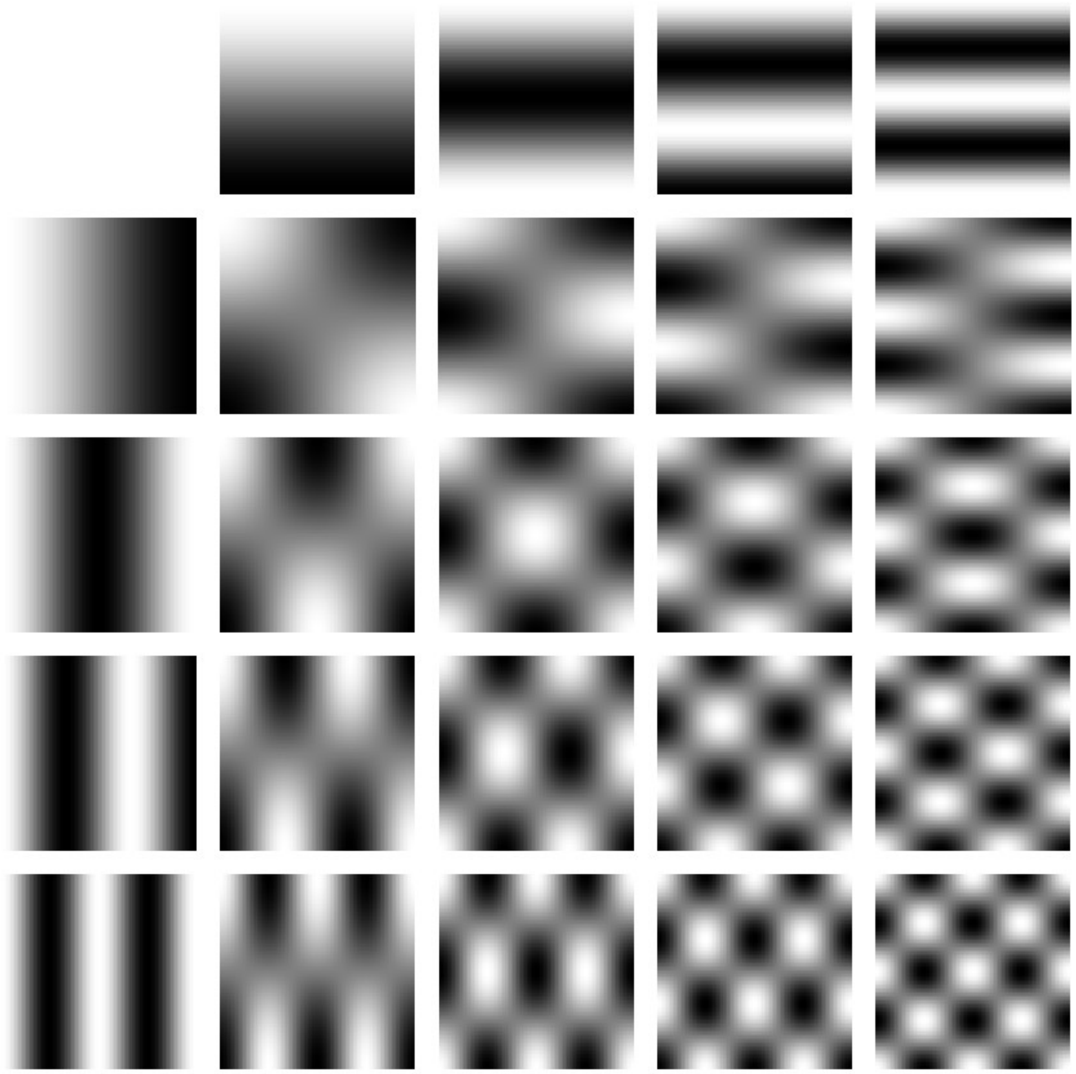
$\phi_m(\mathbf{x})$ are *basis functions*: non-linear, smooth functions of $\mathbf{x}$

Example: cosines (1D)

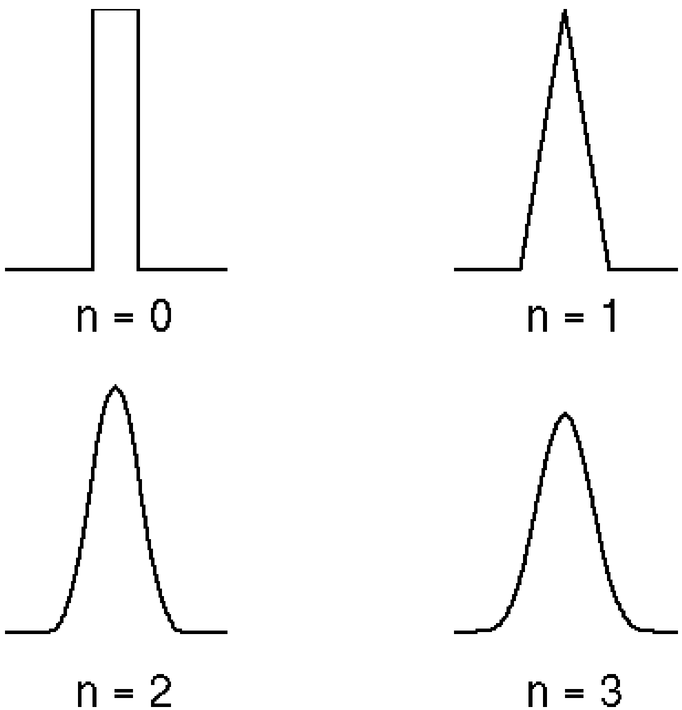


Example: cosines (2D)

$\phi_m(\mathbf{x})$

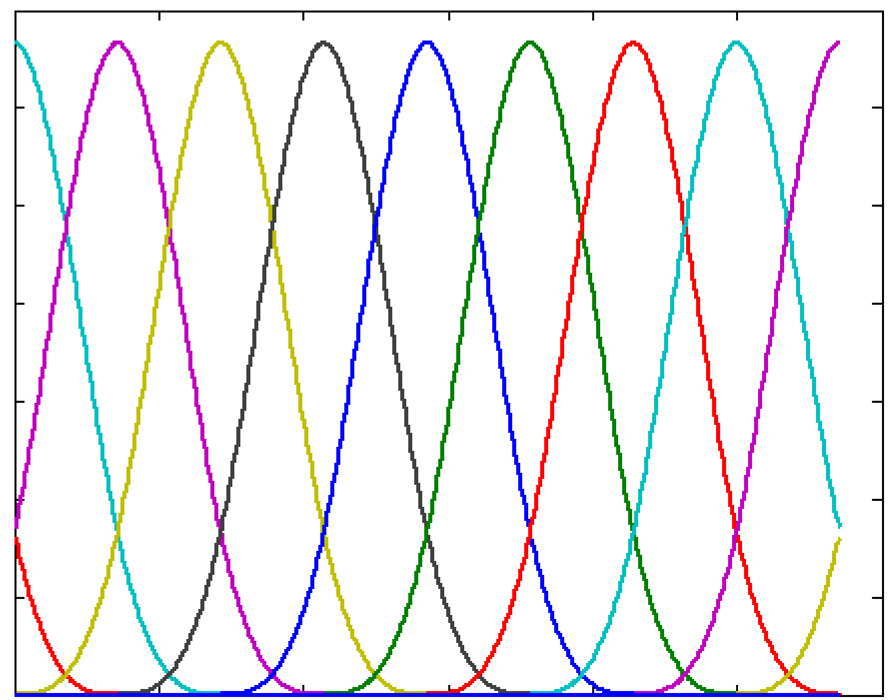


Example: B-splines (1D)



$$B_0(x) = \begin{cases} 1, & -\frac{1}{2} \le x \le \frac{1}{2} \\ \frac{1}{2}, & |x| = \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

$$B_n(x) = \underbrace{B_0 * B_0 \ldots B_0(x)}_{\text{(n+1) times}}.$$



$$\phi_m(\mathbf{x})$$

Linear basis function models

Minimize: $$\sum_{i=1}^{N}(y_i - f(\mathbf{x}_i))^2 = (\mathbf{y} - \mathbf{\Phi}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y} - \mathbf{\Phi}\boldsymbol{\beta})$$

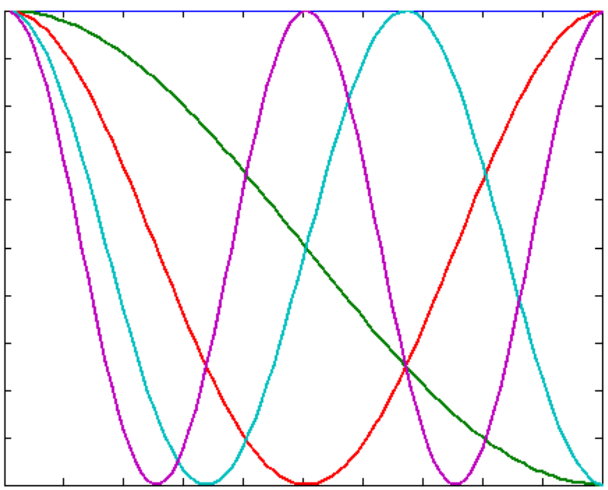
$$\mathbf{\Phi} = \begin{pmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \cdots & \phi_M(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \cdots & \phi_M(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \cdots & \phi_M(\mathbf{x}_N) \end{pmatrix}$$

Non-linear in input variables, but still linear in parameters!

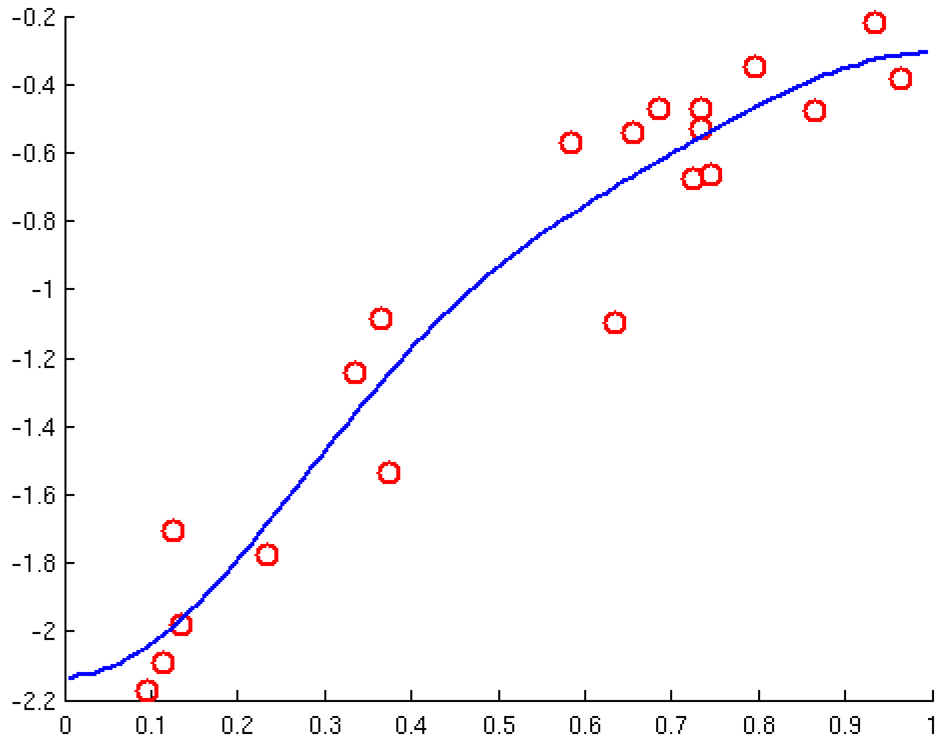
Solution (set gradient to zero): $$\hat{\boldsymbol{\beta}} = (\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi})^{-1}\mathbf{\Phi}^{\mathrm{T}}\mathbf{y}$$

Predicted values at training points: $$\hat{\mathbf{y}} = \mathbf{\Phi}\,\hat{\boldsymbol{\beta}} = \mathbf{\Phi}(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi})^{-1}\mathbf{\Phi}^{\mathrm{T}}\mathbf{y}$$

Linear basis function models



$\phi_m(\mathbf{x})$



Regularization

Minimize: $(\mathbf{y} - \mathbf{\Phi}\boldsymbol{\beta})^{\mathrm{T}}(\mathbf{y} - \mathbf{\Phi}\boldsymbol{\beta}) + \boxed{\lambda\boldsymbol{\beta}^{\mathrm{T}}\mathbf{\Theta}\boldsymbol{\beta}}$

$\mathbf{\Theta}$ is a $M \times M$ positive-semidefinite matrix that penalizes certain $\boldsymbol{\beta}$

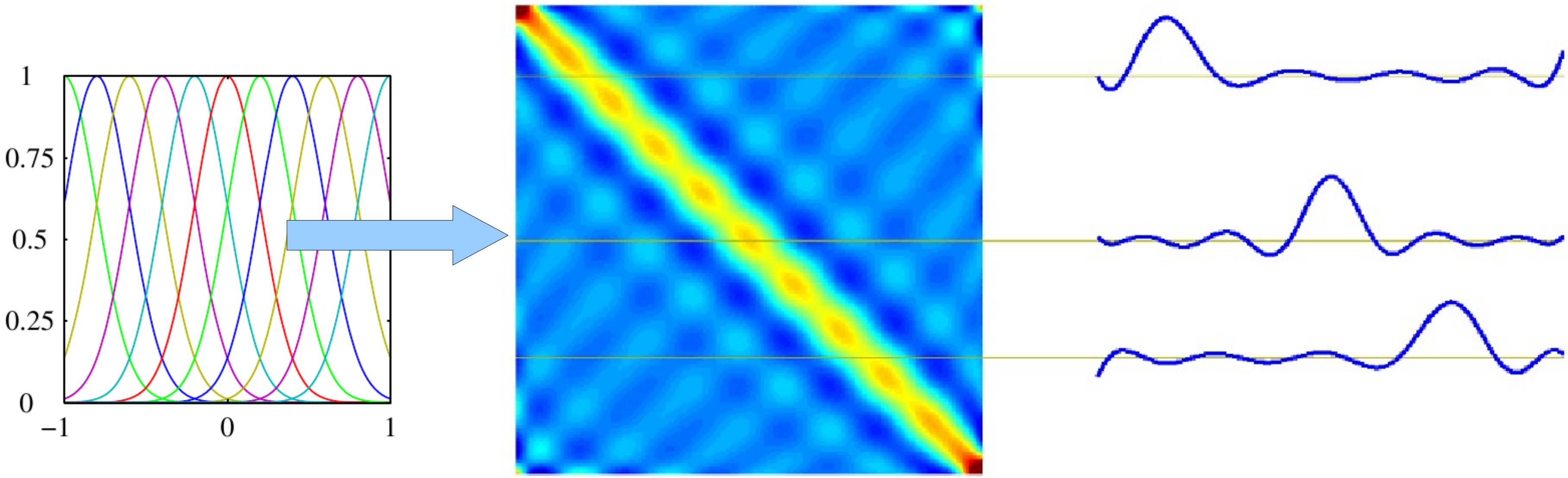
Solution (set gradient to zero): $\hat{\boldsymbol{\beta}} = (\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi} + \lambda\mathbf{\Theta})^{-1}\mathbf{\Phi}^{\mathrm{T}}\mathbf{y}$

Predicted values at training points: $\hat{\mathbf{y}} = \mathbf{\Phi}(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi} + \lambda\mathbf{\Theta})^{-1}\mathbf{\Phi}^{\mathrm{T}}\mathbf{y}$

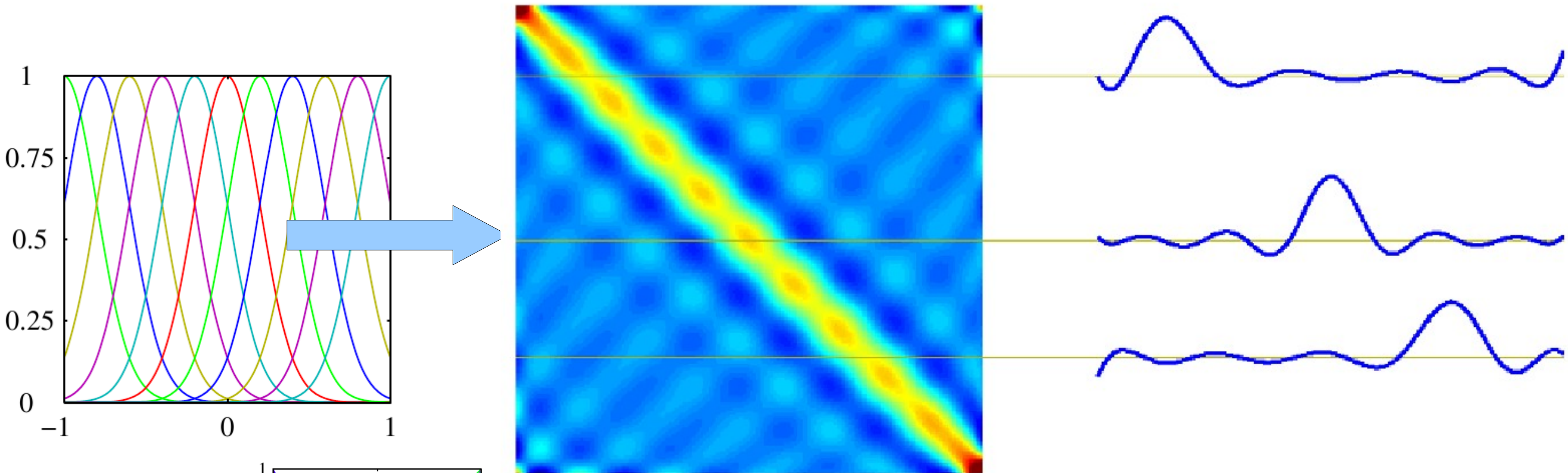
$\mathrm{tr}(\mathbf{\Phi}(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi} + \lambda\mathbf{\Theta})^{-1}\mathbf{\Phi}^{\mathrm{T}})$ $=$ ***effective*** degrees of freedom

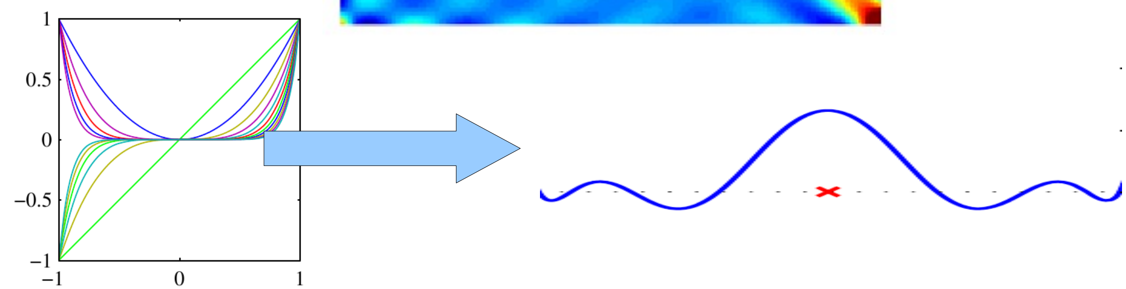
$\leq$ number of parameters M

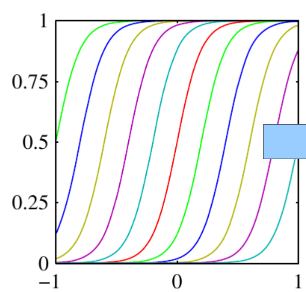
Meaning of $\hat{\mathbf{y}} = \mathbf{\Phi}(\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi} + \lambda\mathbf{\Theta})^{-1}\mathbf{\Phi}^{\mathrm{T}}\mathbf{y}$

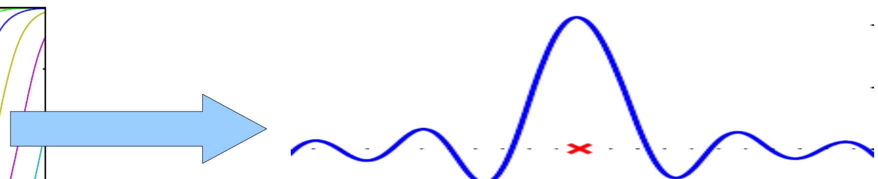


Meaning of $\hat{\mathbf{y}} = \mathbf{\Phi}(\mathbf{\Phi}^T\mathbf{\Phi} + \lambda\mathbf{\Theta})^{-1}\mathbf{\Phi}^T\mathbf{y}$









"buzzGroup_questions.pdf" in DTU Learn (*https://learn.inside.dtu.dk)*

Please discuss in your buzz group

We have seen that approximating a N-dimensional measurement vector $\mathbf{y}$ with a regression model $\mathbf{\Phi}\beta$, where $\mathbf{\Phi}$ is a $N \times M$ matrix of M basis functions stored in its columns, and β is a M-dimensional parameter vector, can be obtained by minimizing

$$\|\mathbf{y} - \mathbf{\Phi}\beta\|^2 + \lambda \beta^T \mathbf{\Theta} \beta,$$

where $\mathbf{\Theta}$ is a regularization matrix and λ is a regularization strength.

Here we will consider the special case where there is no regularization (e.g., λ is set to zero), so that the regression problem reduces to minimizing

$$\|\mathbf{y} - \mathbf{\Phi}\beta\|^2.$$

As we have seen, the solution to this problem is given by

$$\hat{\beta} = \left(\mathbf{\Phi}^T \mathbf{\Phi}\right)^{-1} \mathbf{\Phi}^T \mathbf{y},$$

so that $\mathbf{y}$ is approximated by

$$\hat{\mathbf{y}} = \mathbf{\Phi}\hat{\beta} = \mathbf{\Phi}\left(\mathbf{\Phi}^T \mathbf{\Phi}\right)^{-1} \mathbf{\Phi}^T \mathbf{y} = \mathbf{H}\mathbf{y},$$

where $\mathbf{H} = \mathbf{\Phi}\left(\mathbf{\Phi}^T \mathbf{\Phi}\right)^{-1} \mathbf{\Phi}^T$ is a $N \times N$ matrix that transforms $\mathbf{y}$ into $\hat{\mathbf{y}}$.

Task 1

Consider the case

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{\Phi} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

What is $\hat{\beta}$?

Task 2

In the same case

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{\Phi} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

what is $\mathbf{H}$, what is the number of degrees of freedom $\mathrm{trace}(\mathbf{H})$, and what is $\hat{\mathbf{y}}$? Can you explain (e.g., draw) what is happening?

Task 3

Now consider the slightly harder problem

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{\Phi} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}.$$

Again: what is $\mathbf{H}$, what is the number of degrees of freedom $\mathrm{trace}(\mathbf{H})$, and what is $\hat{\mathbf{y}}$? Can you explain (e.g., draw) what is happening?

Solution to task 1

Consider the case

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Phi} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

What is $\hat{\beta}$?

Solution to task 1

Consider the case

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Phi} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

What is $\hat{\beta}$?

Solution:

$$\boldsymbol{\Phi}^T \mathbf{y} = 1 + 2 + 3 = 6,$$

and

$$\boldsymbol{\Phi}^T \boldsymbol{\Phi} = 1^2 + 1^2 + 1^2 = 3,$$

so

$$\hat{\beta} = \left(\boldsymbol{\Phi}^T \boldsymbol{\Phi}\right)^{-1} \boldsymbol{\Phi}^T \mathbf{y} = 3^{-1} 6 = 2.$$

Solution to task 2

In the same case

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Phi} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

what is $\mathbf{H}$, what is the number of degrees of freedom $\mathrm{tracc}(\mathbf{H})$, and what is $\hat{\mathbf{y}}$? Can you explain (e.g., draw) what is happening?

Solution to task 2

In the same case

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{\Phi} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

what is $\mathbf{H}$, what is the number of degrees of freedom $\mathrm{trace}(\mathbf{H})$, and what is $\hat{\mathbf{y}}$? Can you explain (e.g., draw) what is happening?

Solution:

$$\begin{aligned}
\mathbf{H} &= \mathbf{\Phi}\left(\mathbf{\Phi}^T\mathbf{\Phi}\right)^{-1}\mathbf{\Phi}^T \\
&= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot 3^{-1} \cdot \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \\
&= 3^{-1} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \\
&= 3^{-1} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix},
\end{aligned}$$

so

$$\mathrm{trace}(\mathbf{H}) = 1/3 + 1/3 + 1/3 = 1,$$

and

$$\hat{\mathbf{y}} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

Interpretation: We are fitting a horizontal line to three points. Offset is the only parameter we get to play with (one degree of freedom). Figure 1 illustrates the situation.

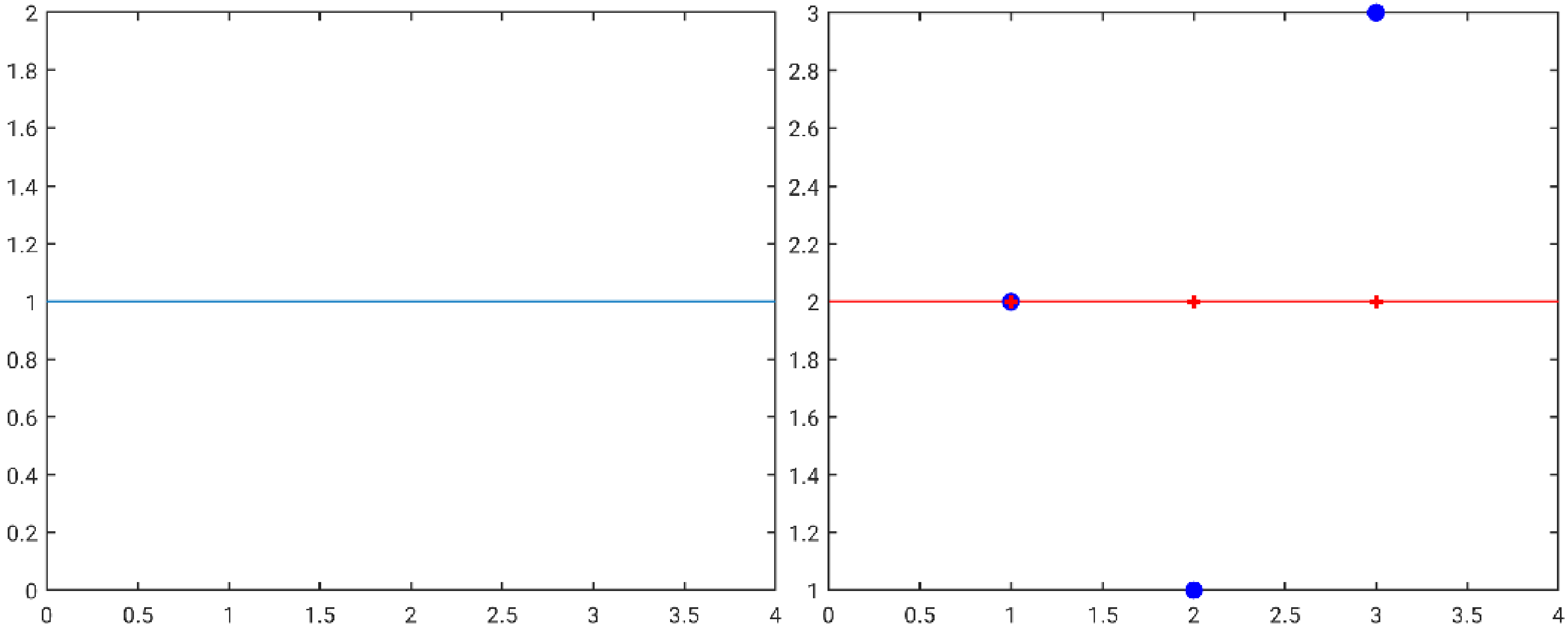


Figure 1: Left: the one column in $\mathbf{\Phi}$ represents a horizontal line. Right: the horizontal line is fitted to the input measurements $\mathbf{y}$ (blue dots), resulting in $\hat{\mathbf{y}}$ (red crosses).

Solution to task 3

Now consider the slightly harder problem

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{\Phi} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}.$$

Again: what is $\mathbf{H}$, what is the number of degrees of freedom $\mathrm{trace}(\mathbf{H})$, and what is $\hat{\mathbf{y}}$? Can you explain (e.g., draw) what is happening?

Solution to task 3

Now consider the slightly harder problem

$$\mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Phi} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}.$$

Again: what is $\mathbf{H}$, what is the number of degrees of freedom $\mathrm{trace}(\mathbf{H})$, and what is $\hat{\mathbf{y}}$? Can you explain (e.g., draw) what is happening?

Solution: since $\boldsymbol{\Phi}$ is square and invertible, we have that

$$\left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1} = \boldsymbol{\Phi}^{-1}\left(\boldsymbol{\Phi}^T\right)^{-1},$$

so that

$$\begin{aligned}
\mathbf{H} &= \boldsymbol{\Phi}\left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^T \\
&= \overbrace{\boldsymbol{\Phi}\boldsymbol{\Phi}^{-1}}^{\mathbf{I}}\underbrace{\left(\boldsymbol{\Phi}^T\right)^{-1}\boldsymbol{\Phi}^T}_{\mathbf{I}} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}$$

Therefore

$$\mathrm{trace}(\mathbf{H}) = 1 + 1 + 1 = 3,$$

and $\hat{\mathbf{y}}$ is simply $\mathbf{y}$:

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{y} = \mathbf{y} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

Interpretation: we're fitting three basis functions (quadratic function) to three measurement points. We have three coefficients to play with, so three degrees of freedom for three measurement points, and as a result we can fit the curve perfectly through the measurement points (no smoothing). Figure 2 illustrates the situation.

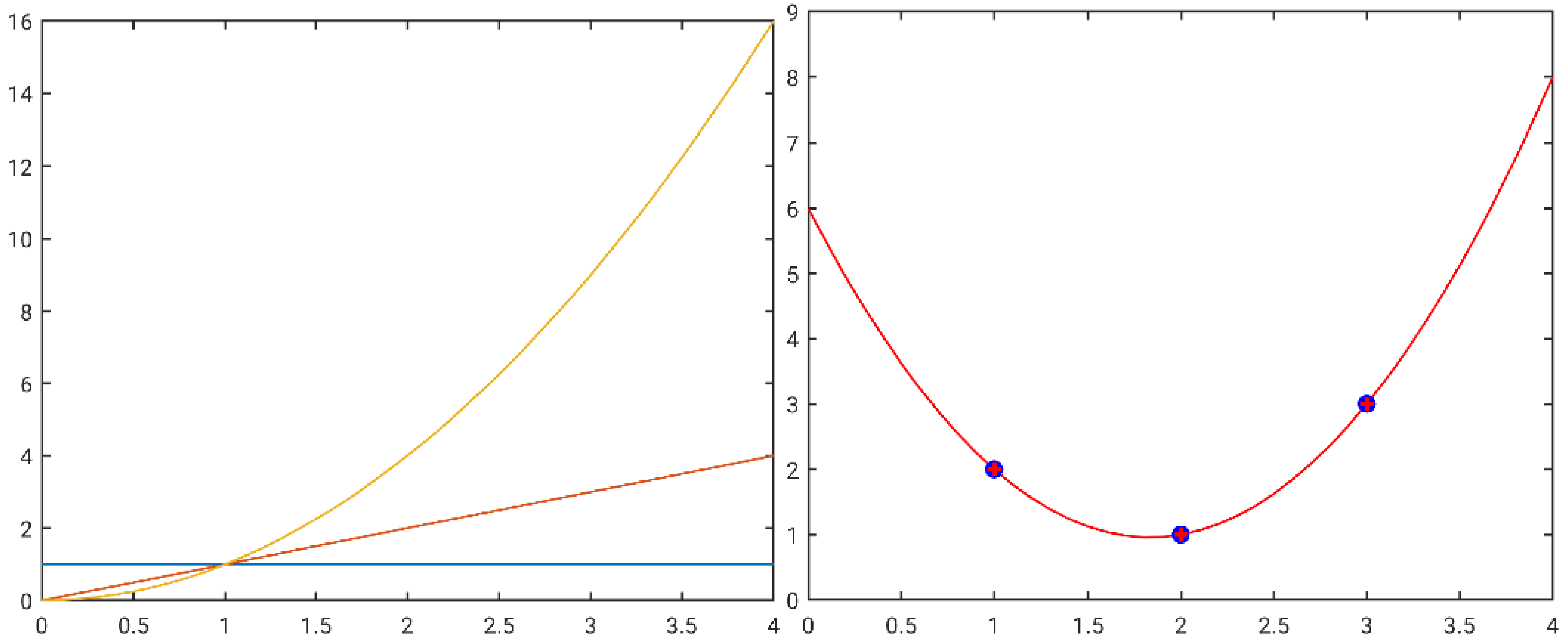


Figure 2: Left: the three columns in Φ represents polynomials of degrees 0, 1 and 2. Right: a quadratic curve is fitted to the input measurements $\mathbf{y}$ (blue dots), resulting in $\hat{\mathbf{y}}$ (red crosses).

Cubic smoothing spline

So far, ***imposed*** smoothness by limiting models to be linear combinations of smooth basis functions

What happens if we don't impose such structure, but simply look for the function $f(x)$ that minimizes

$$\sum_{i=1}^{N}(y_i - f(x_i))^2 + \lambda \int (f''(t))^2 dt$$? (1D)

Cubic smoothing spline

So far, ***imposed*** smoothness by limiting models to be linear combinations of smooth basis functions

What happens if we don't impose such structure, but simply look for the function $f(x)$ that minimizes

$$\sum_{i=1}^{N}(y_i - f(x_i))^2 + \lambda \int (f''(t))^2 dt \quad \text{? (1D)}$$

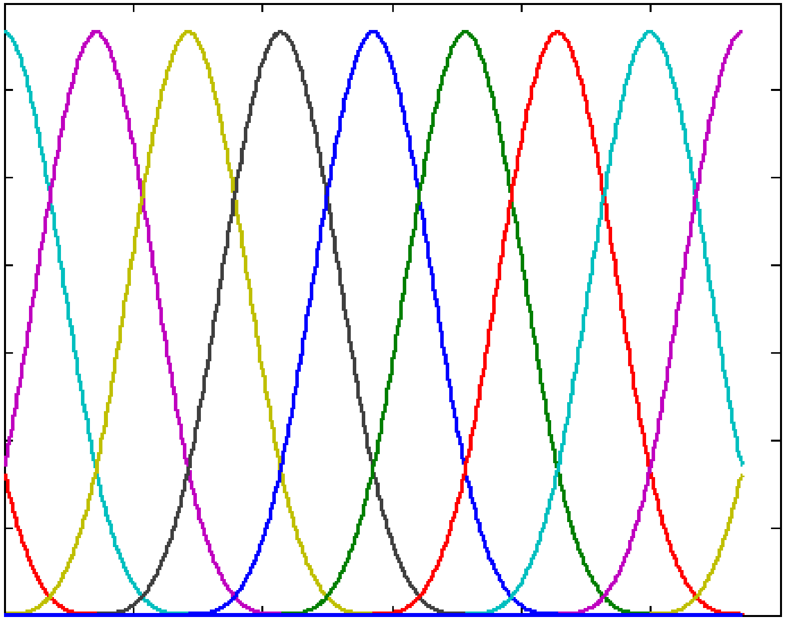
It can be shown to have the structure

$$f(x) = \sum_{m=1}^{M} \beta_m \phi_m(x) \quad !$$

Cubic smoothing spline

$\phi_m(x)$ can be shown to be so-called cubic B-spline basis functions $B_m(x;\xi)$

for equidistantly placed input points x_i, these are the uniform cubic B-spline basis functions encountered before



Cubic smoothing spline

Plugging $f(x) = \sum_{m=1}^{M} \beta_m \phi_m(x)$

into $\sum_{i=1}^{N} (y_i - f(x_i))^2 + \lambda \int (f''(t))^2 dt$,

we have to minimize $(\mathbf{y} - \mathbf{B}\boldsymbol{\beta})^{\mathrm{T}} (\mathbf{y} - \mathbf{B}\boldsymbol{\beta}) + \lambda \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{\Omega}_{\mathbf{B}} \boldsymbol{\beta}$

$\mathbf{B}$ contains the basis functions

$\{\boldsymbol{\Omega}_{\mathbf{B}}\}_{jk} = \int B_j''(t;\xi) B_k''(t;\xi) dt$

Cubic smoothing spline

We've seen this before!

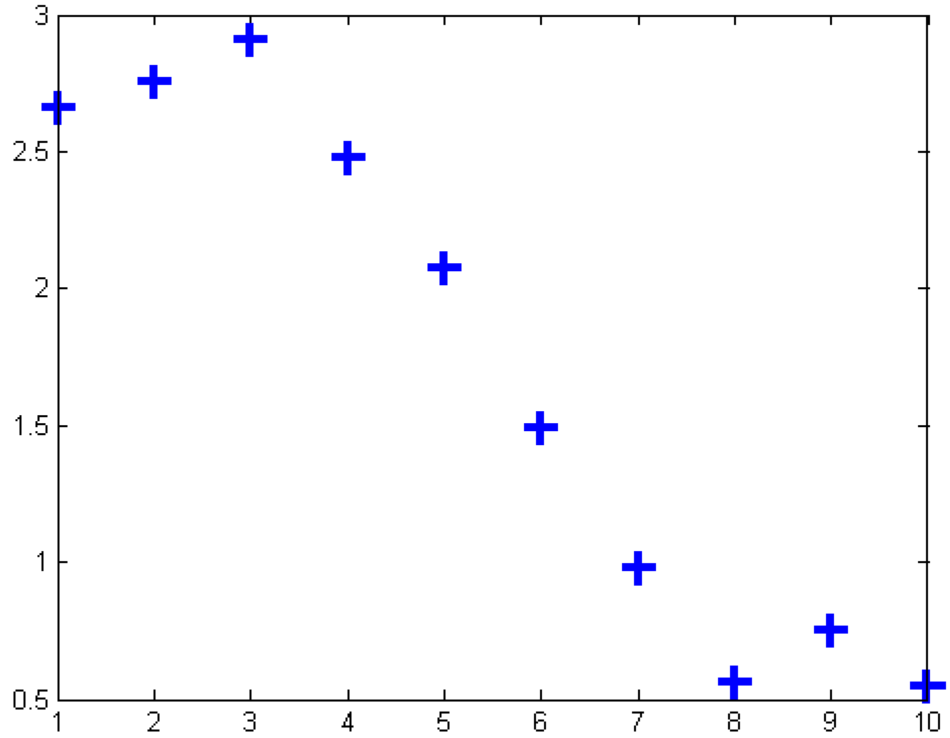
$$\hat{\beta} = (\mathbf{B}^T\mathbf{B} + \lambda\Omega_{\mathbf{B}})^{-1}\mathbf{B}^T\mathbf{y}$$

Predicted values at training points:

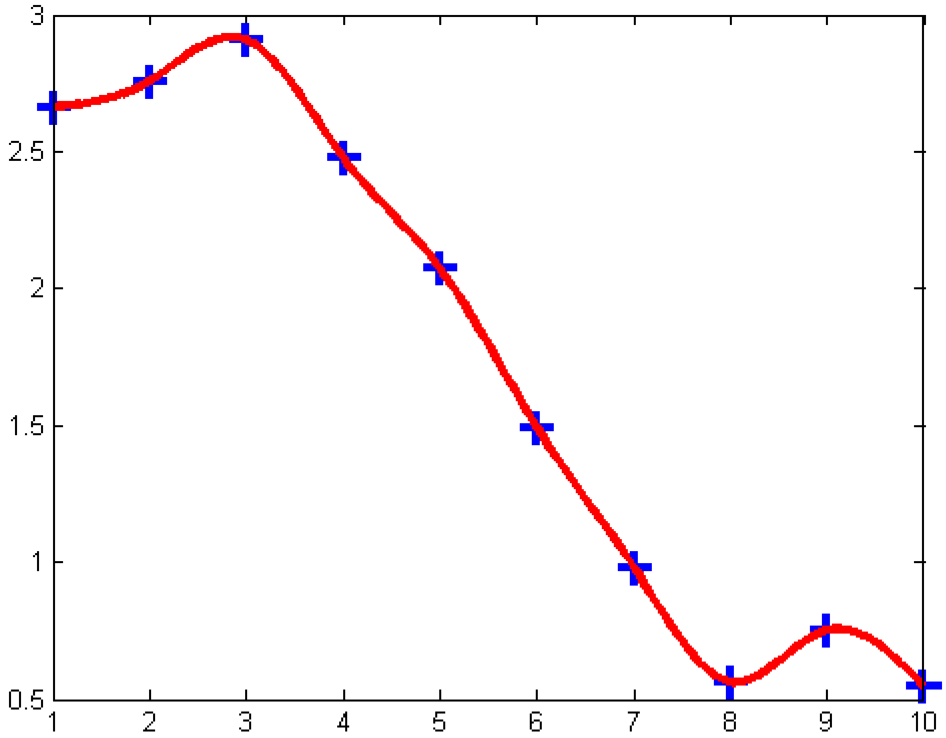
$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{B}(\mathbf{B}^T\mathbf{B} + \lambda\Omega_{\mathbf{B}})^{-1}\mathbf{B}^T\mathbf{y} \\ &= \mathbf{S}_\lambda\mathbf{y}.\end{aligned}$$

Effective degrees of freedom, $\mathrm{tr}(\mathbf{S}_\lambda)$, depends on λ

Cubic smoothing spline

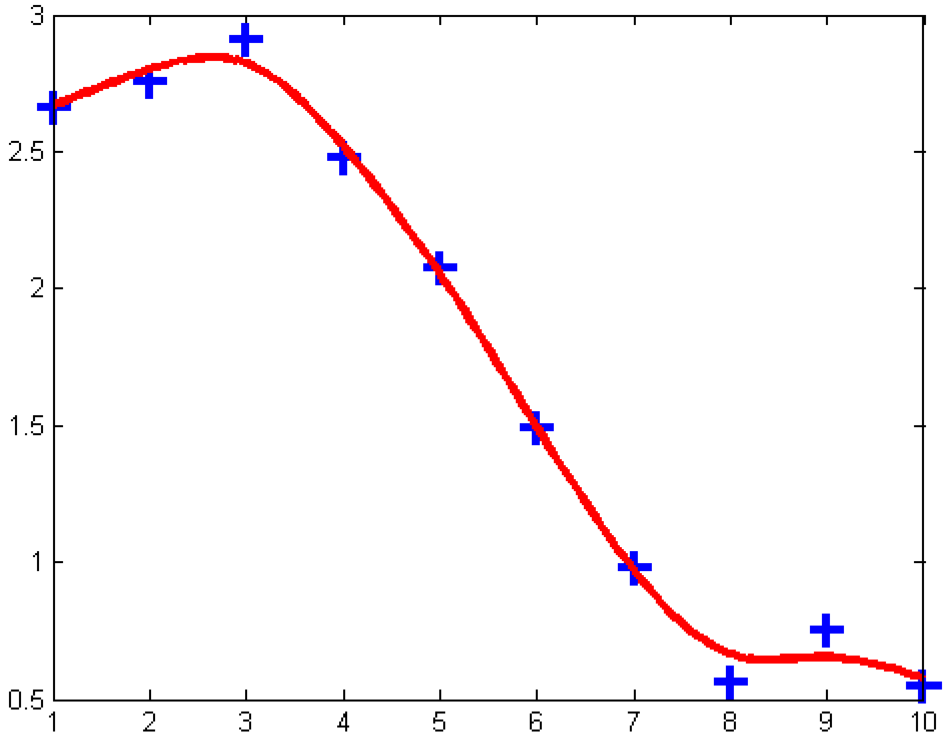


Cubic smoothing spline



$\lambda = 0$: spline passes exactly through measurements

Cubic smoothing spline



$\lambda > 0$: spline smooths measurements

Thin plate spline

Generalization of cubic smoothing spline to 2D and 3D

Find function $f(\mathbf{x})$ that minimizes $\sum_{i=1}^{N}(y_i - f(\mathbf{x}_i))^2 + \lambda J_p(f)$

with $J_p(f) = \int_{\mathbb{R}^p} \sum_{i=1}^{p} \sum_{j=1}^{p} \left[\frac{\partial^2 f}{\partial u_i \partial u_j}\right]^2 d\mathbf{u}$

e.g., in 2D: $J_2(f) = \int_{\mathbb{R}^2} \left[\frac{\partial^2 f}{\partial u_1^2}\right]^2 + 2\left[\frac{\partial^2 f}{\partial u_1 \partial u_2}\right]^2 + \left[\frac{\partial^2 f}{\partial u_2^2}\right]^2 d\mathbf{u}$

Thin plate spline

Can be shown to be of the form

$$f(\mathbf{x}) = \beta_0 + \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x} + \sum_{i=1}^{N} \alpha_i \eta_p(\|\mathbf{x} - \mathbf{x}_i\|),$$

with

$$\eta_2(r) = \begin{cases} r^2 \log r & \text{if } r > 0 \\ 0 & \text{if } r = 0 \end{cases} \qquad \text{(2D)}$$

$$\eta_3(r) = \|r\|^3. \qquad \text{(3D)}$$

Thin plate spline

Plugging $$f(\mathbf{x}) = \beta_0 + \boldsymbol{\beta}_1^{\mathrm{T}}\mathbf{x} + \sum_{i=1}^{N} \alpha_i \eta_p(\|\mathbf{x} - \mathbf{x}_i\|),$$

into $$\sum_{i=1}^{N} (y_i - f(\mathbf{x}_i))^2 + \lambda J_p(f)$$

yields minimization problem of the form (see course notes):

$$\begin{aligned} \|\mathbf{y} - \mathbf{P}^{\mathrm{T}}\boldsymbol{\beta} - \mathbf{K}\boldsymbol{\alpha}\|^2 + \lambda\boldsymbol{\alpha}^{\mathrm{T}}\mathbf{K}\boldsymbol{\alpha} \\ \text{s.t.} \quad \mathbf{P}\boldsymbol{\alpha} \quad = \quad 0 \end{aligned}$$

Solution (see course notes):

$$\begin{pmatrix} \mathbf{K} + \lambda\mathbf{I} & \mathbf{P}^{\mathrm{T}} \\ \mathbf{P} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}$$

Thin plate spline

Predicted values at training points:

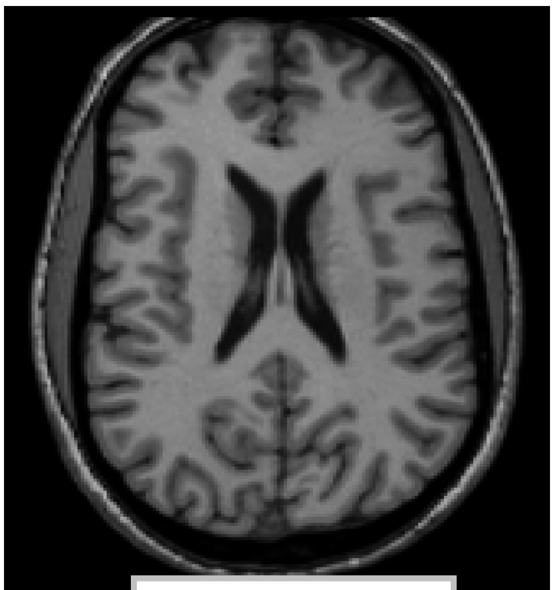
$$\hat{\mathbf{y}} = \begin{pmatrix} \mathbf{K} & \mathbf{P}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \mathbf{K} & \mathbf{P}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \mathbf{K} + \lambda\mathbf{I} & \mathbf{P}^{\mathrm{T}} \\ \mathbf{P} & \mathbf{0} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix} = \mathbf{S}_{\lambda}^{*} \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}$$

$N \times N$ "smoother matrix" $\mathbf{S}_{\lambda}$ is the first N columns of $\mathbf{S}_{\lambda}^{*}$

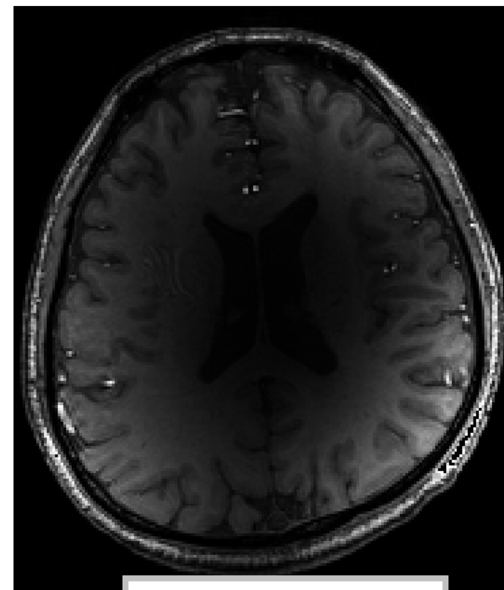
Effective degrees of freedom is again $\mathrm{trace}(\mathbf{S}_{\lambda})$

$\lambda = 0$: spline passes exactly through measurements
("thin plate spline")

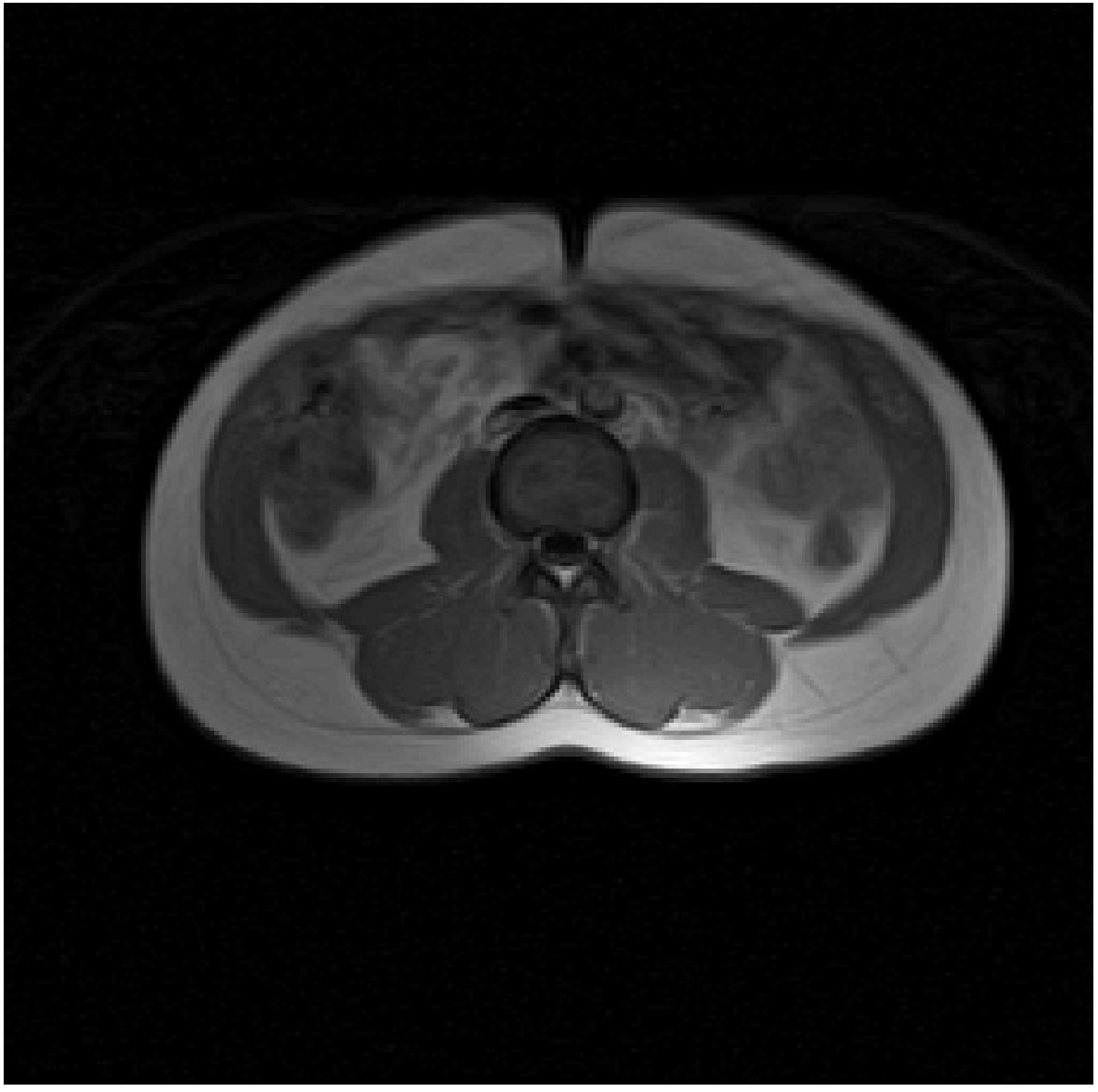
Application: MR bias field


1.5 Tesla scanner

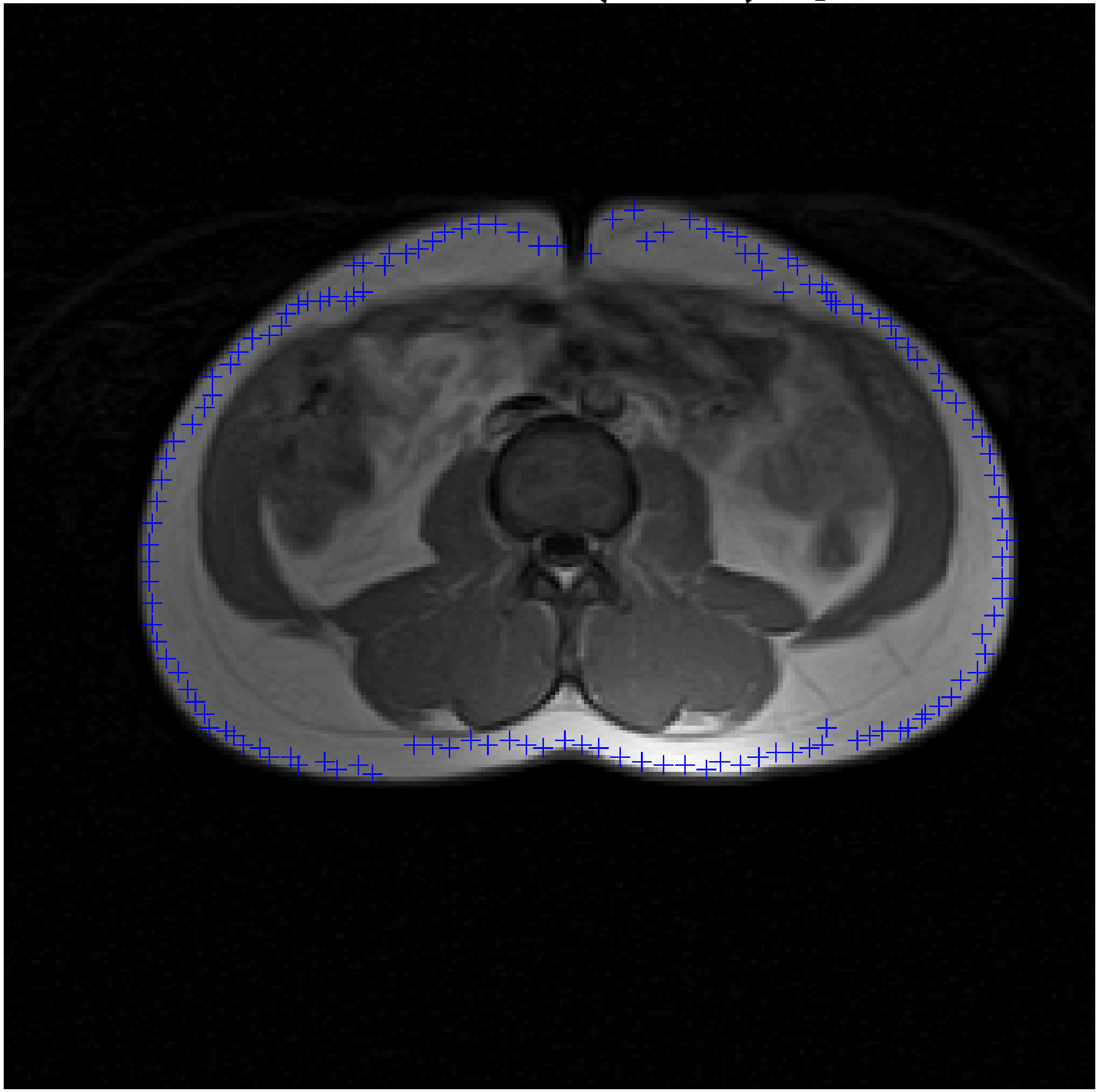

3 Tesla scanner


7 Tesla scanner

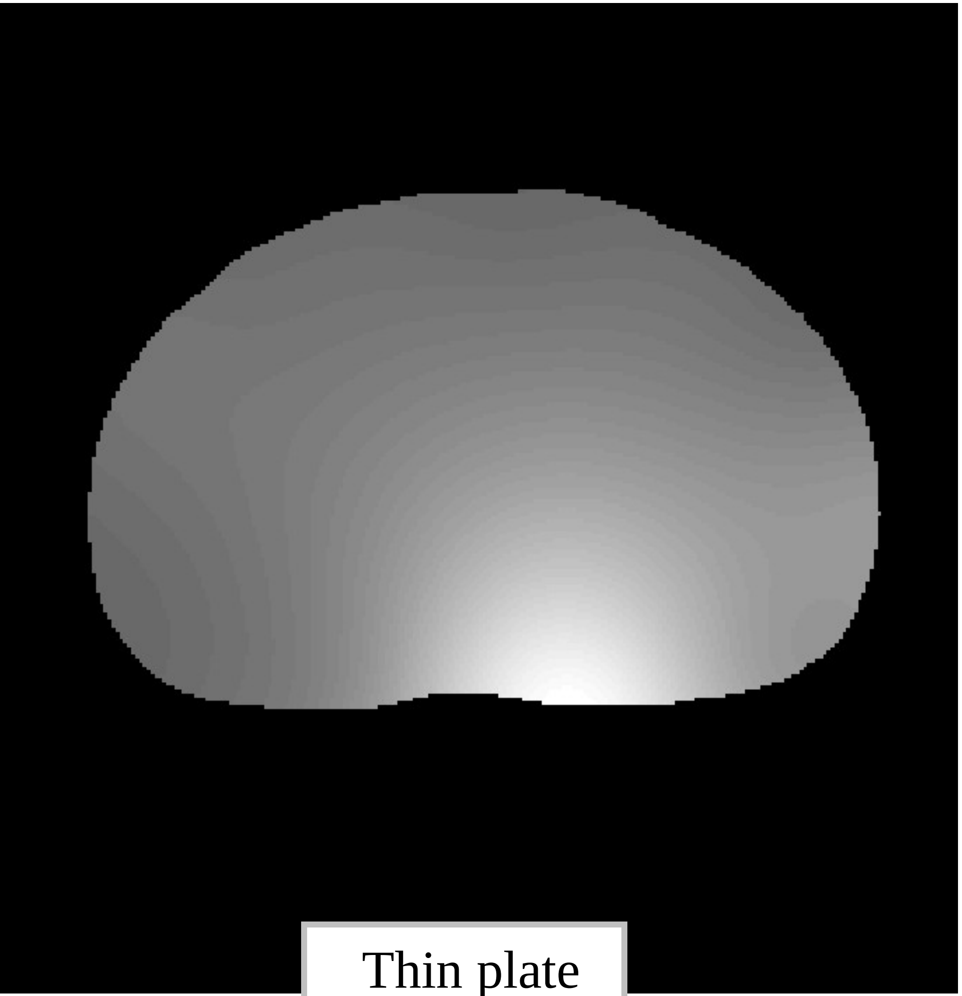
Abdominal MR slice



Same tissue (fat) pixels



Bias and corrected image



Thin plate spline