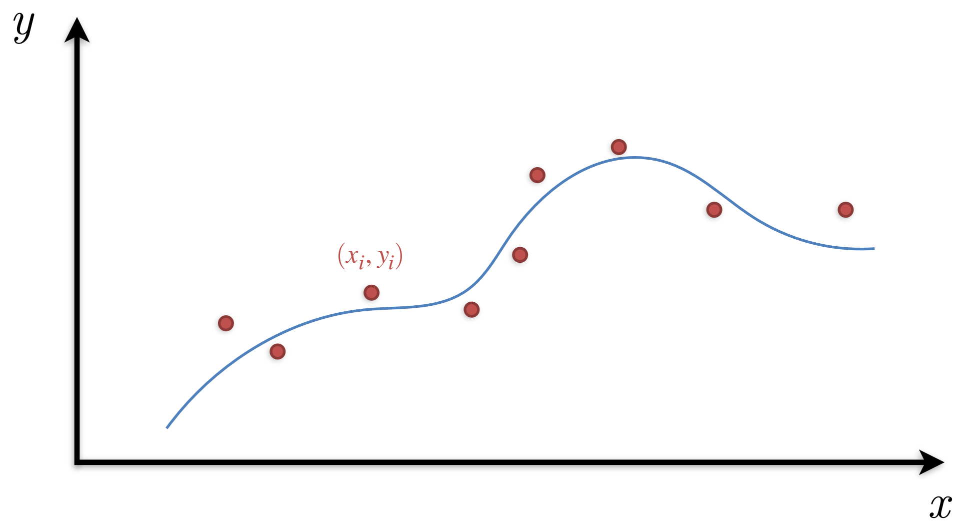
EPFL

Linear Least Squares EE-312

Prof. Pierre Vandergheynst

Motivation

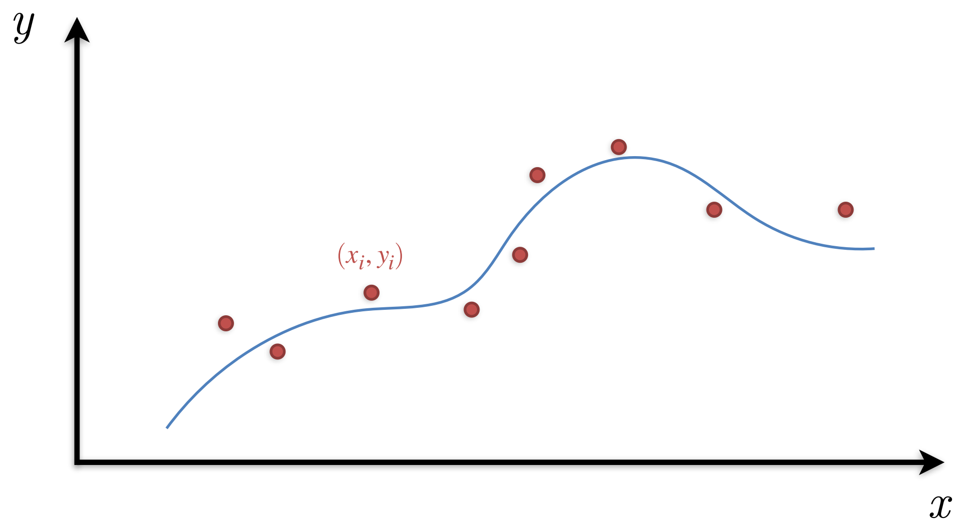


$y_i,\ i = 1, \ldots, m$ measurements taken at corresponding x_i

Hypothesis: $y = f(x; \beta) = \sum_{j=1}^{n} \beta_j \varphi_j(x)$

A linear model (in the unknown parameters β_j)

Motivation



$$y = Ax + \delta$$

It is unlikely we achieve $y_i = f(x_i;\beta)\,\forall i \Rightarrow y_i = f(x_i;\beta) + \delta_i$ Small error

Set $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m \quad A = \begin{pmatrix} \varphi_1(y_1) & \dots & \varphi_n(y_1) \\ \vdots & & \vdots \\ \varphi_1(y_m) & \dots & \varphi_n(y_m) \end{pmatrix} \in \mathbb{R}^{m\times n} \quad x = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{R}^n \quad \delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_m \end{pmatrix} \in \mathbb{R}^m$

Motivation

With these notations, we are led to looking for the coefficients that will minimise the error:

$$\arg\min_{x\in\mathbb{R}^n} \|Ax - y\|_2$$

Before proceeding, let us informally explore ...

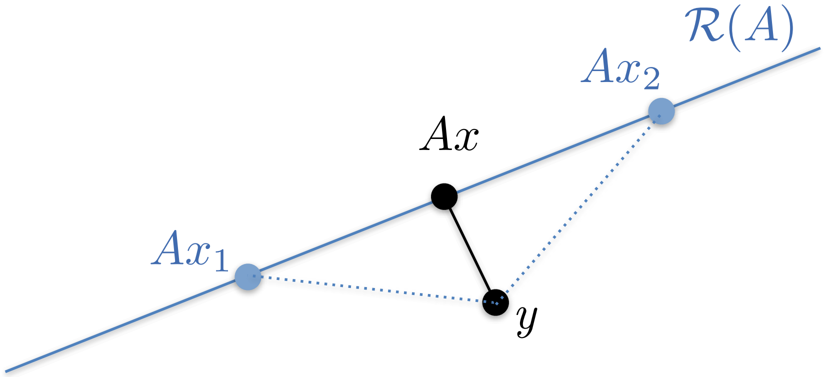
$$\nabla\left(\|Ax - y\|_2^2\right) = 2A^T\left(Ax - y\right)$$

Setting the gradient to zero, we find something interesting:

$$A^T Ax = A^T y \Rightarrow x = \left(A^T A\right)^{-1} A^T y$$

Motivation

Now let's try geometrically:



$$(y - Ax) \perp \mathcal{R}(A) \quad \text{at a minimiser}$$

$$(Au)^T(y - Ax) = 0 \text{ for arbitrary } u \in \mathbb{R}^n$$

$$u^T A^T (y - Ax) = 0$$

$$u^T (A^T y - A^T Ax) = 0$$

And again, we reach the same characterisation:

$$A^T Ax = A^T y \Rightarrow x = \left(A^T A\right)^{-1} A^T y$$

Linear Least Squares

We will study the set of minimisers of our problem:

$$\mathcal{X} = \{x \in \mathbb{R}^n \text{ s.t. } \rho(x) = \|Ax - y\|_2^2 \text{ is minimized}\}$$

Normal equations: $x \in \mathcal{X}$ IFF $A^T r = 0$ where the residual $r = (y\text{-}Ax)$

The normal equations can be rewritten as: $A^T A x = A^T y$

Characterisation of solutions: $x \in \mathcal{X}$ IFF it is of the form

$$x = A^+ y + \left(\mathcal{I} - A^+ A\right) b, \text{ where } b \in \mathbb{R}^n \text{ is arbitrary}$$

Moreover $x \in \mathcal{X} \Rightarrow \rho(x) \leqslant \|y\|_2^2$

Linear Least Squares

We will study the set of minimisers of our problem:

$$\mathcal{X} = \{x \in \mathbb{R}^n \text{ s.t. } \rho(x) = \|Ax - y\|_2^2 \text{ is minimized}\}$$

The set $\mathcal{X}$ is convex

$$x_1, x_2 \in \mathcal{X},\ \theta \in [0,1] \Rightarrow \theta x_1 + (1-\theta)x_2 \in \mathcal{X}$$

Unicity: there is a unique solution, i.e $\mathcal{X} = \{x^*\}$

IFF $A^+A = \mathbb{I}$ or equivalently $\text{rank}(A) = n$

In that case $x^* = A^+y$

Linear Least Squares

We will study the set of minimisers of our problem:

$$\mathcal{X} = \{x \in \mathbb{R}^n \text{ s.t. } \rho(x) = \|Ax - y\|_2^2 \text{ is minimized}\}$$

General case: rank deficient matrix A

There is a unique $x^* \in \mathcal{X}$ of minimum 2-norm

$$x^* = A^+ y$$

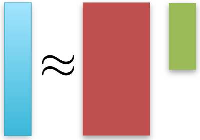
A General View

Solve: $y = Ax,\ A \in \mathbb{R}^{m \times n}$

Determined (m=n):

Over-Determined (m>n): $\arg\min_{x \in \mathbb{R}^n} \|Ax - y\|_2$

$$x = \left(A^T A\right)^{-1} A^T y = A^+ y$$



Under-Determined (m<n): $\arg\min_{x \in \mathbb{R}^n} \{\|x\|_2 \text{ such that } Ax = y\}$

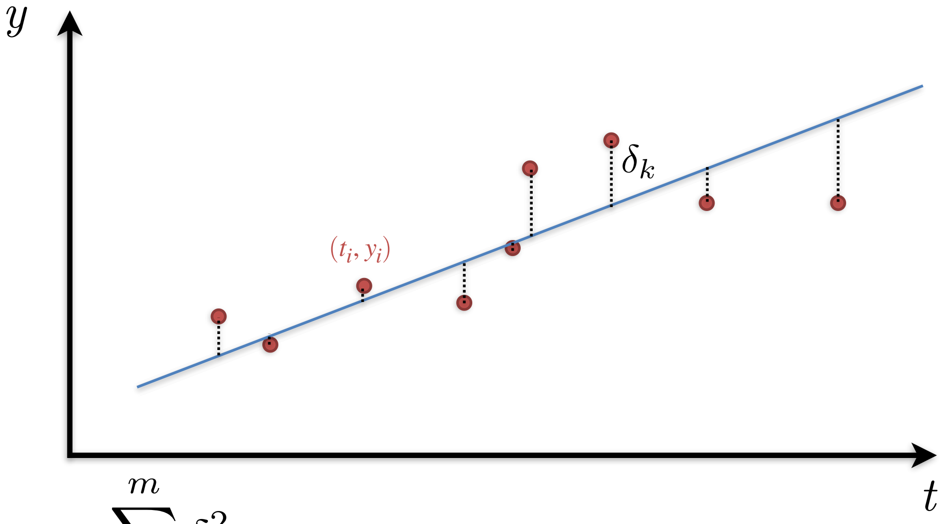
$$x = A^T \left(AA^T\right)^{-1} y = A^+ y$$

Simple Linear Regression

m measurements: $(t_1, y_1), \cdots, (t_m, y_m)$

Looking for a simple linear model: $y = \alpha t + \beta$, with $y, t \in \mathbb{R}^m$ and $\alpha, \beta \in \mathbb{R}$

2 parameters & m measurements (over-determined)



$$y_1 = \alpha t_1 + \beta + \delta_1$$

$$\vdots$$

$$y_m = \alpha t_m + \beta + \delta_m$$

Minimize the sum of squared errors $\sum_{i=1}^{m} \delta_i^2$

Simple Linear Regression

m measurements: $(t_1, y_1), \cdots, (t_m, y_m)$

Looking for a simple linear model: $y = \alpha t + \beta,$ with $y, t \in \mathbb{R}^m$ and $\alpha, \beta \in \mathbb{R}$

Minimize the sum of squared errors $\sum_{i=1}^{m} \delta_i^2$

$$y_1 = \alpha t_1 + \beta + \delta_1$$
$$\vdots$$
$$y_m = \alpha t_m + \beta + \delta_m$$

In matrix form: $y = Ax + \delta$ $\quad \arg\min_x \|y - Ax\|_2^2$

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad A = \begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix} \quad x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_m \end{pmatrix}$$

Simple Linear Regression

m measurements: $(t_1, y_1), \cdots, (t_m, y_m)$

Looking for a simple linear model: $y = \alpha t + \beta,$ with $y, t \in \mathbb{R}^m$ and $\alpha, \beta \in \mathbb{R}$

$$\arg\min_x \|y - Ax\|_2^2$$

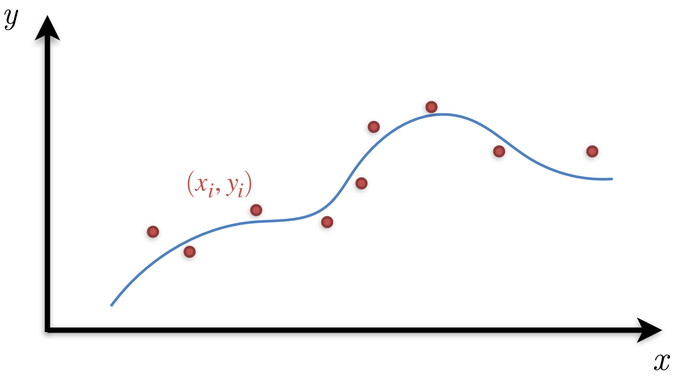
$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \quad A = \begin{pmatrix} t_1 & 1 \\ \vdots & \vdots \\ t_m & 1 \end{pmatrix} \quad x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_m \end{pmatrix}$$

Normal equations: $A^T A x = A^T y$

A full rank (at least two distinct measurements to fit our line)

$$A^T A x = A^T y \Rightarrow x = \left(A^T A\right)^{-1} A^T y$$

Less Simple Linear Regression ?



Hypothesis: $y = f(x;\beta) = \sum_{j=1}^{n} \beta_j \varphi_j(x)$ A linear model (in the unknown parameters β_j)

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m \quad A = \begin{pmatrix} \varphi_1(y_1) & \dots & \varphi_n(y_1) \\ \vdots & & \vdots \\ \varphi_1(y_m) & \dots & \varphi_n(y_m) \end{pmatrix} \in \mathbb{R}^{m\times n} \quad x = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{R}^n$$