

ETH Lecture 401-0663-00L Numerical Methods for CSE

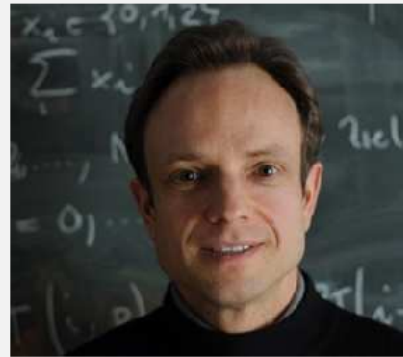
Course Video

Section 5.7: Least Squares Data Fitting

Prof. R. Hiptmair, SAM, ETH Zurich

Date: November 8, 2020

(C) Seminar für Angewandte Mathematik, ETH Zürich



Dependencies. Section 3.1, Section 3.6.1, Section 5.2

Duration: 26 minutes

V. Data Interpolation and Data Fitting in 1D

5.7. Least-Squares Data Fitting

Setting: Data point $(t_i, y_i) \in \mathbb{R}^2$, $i = 0, \dots, m$, $t_i \in I$,
affected measurement errors/uncertain

▷ Interpolation dubious



replace with least-squares fitting

$V \cong$ vector space of admissible functions, $V \subset C^0(I)$

$$f \in \underset{g \in V}{\operatorname{argmin}} \sum_{i=0}^m |g(t_i) - y_i|^2 \quad (*)$$

Example: polynomial fitting $V = \mathcal{P}_p$, $p \leq m$

Assume: $V = \text{Span} \underbrace{\{b_1, \dots, b_n\}}_{\text{basis}}, n \leq m$

$$(*) \Leftrightarrow g = \sum_{j=1}^n (x)_j b_j$$

$$x \in \underset{z \in \mathbb{R}^n}{\text{argmin}} \underbrace{\sum_{i=0}^m \left| \sum_{j=1}^n (z)_j b_j(t_i) - y_i \right|^2}_{= \|Az - y\|_2^2}$$

with $A = [b_j(t_i)]_{\substack{i=0, \dots, m \\ j=1, \dots, n}} \in \mathbb{R}^{m+1, n}$

$$y = [y_0, \dots, y_m] \in \mathbb{R}^{m+1}$$

$\hat{=}$ A std. linear least squares problem

$$x \text{ unique} \Leftrightarrow \text{rank}(A) = n$$

Lemma 5.7.0.17. Unique solvability of linear least squares fitting problem

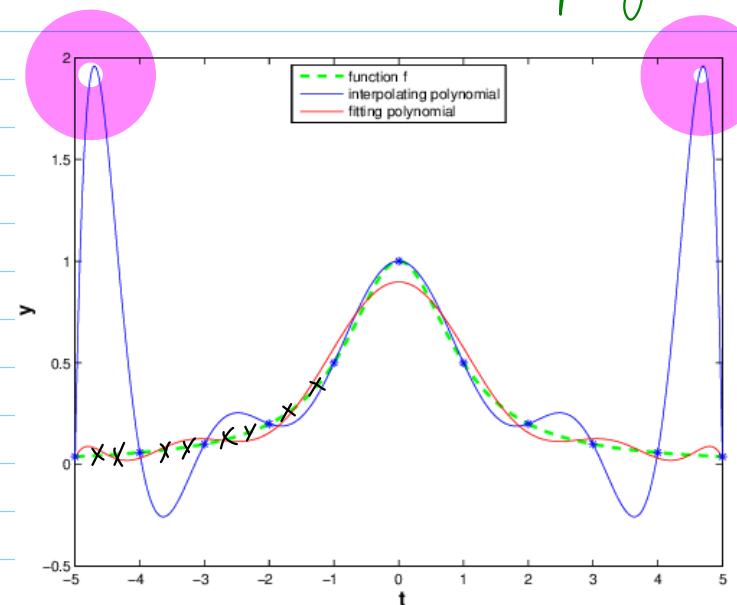
The scalar one-dimensional linear least squares fitting problem (5.7.0.10) with $\dim V_n = n$, V_n the vector space of admissible functions, has a unique solution, if and only if there are $t_{i_1}, \dots, t_{i_n}$ such that

$$\begin{bmatrix} b_1(t_{i_1}) & \dots & b_n(t_{i_1}) \\ \vdots & & \vdots \\ b_1(t_{i_n}) & \dots & b_n(t_{i_n}) \end{bmatrix} \in \mathbb{R}^{n,n} \text{ is invertible,} \quad (5.7.0.18)$$

which is *independent* of the choice of *basis* of V_n .

Existence of an n -subset of nodes for which interpolation into V will always have a unique solution

Exp. 5.7.0.23 (Polynomial interpolation vs. polynomial fitting)



Data from function $f(t) = \frac{1}{1+t^2}$,

- polynomial degree $d = 10$,
- interpolation through data points $(t_j, f(t_j))$, $j = 0, \dots, d$, $t_j = -5 + j$, see Ex. 5.2.4.3,
- fitting to data points $(t_j, f(t_j))$, $j = 0, \dots, 3d$, $t_j = -5 + j/3$.

▷ Better "shape fidelity" by fitting.

③

Rem 5.7.0.24 (Overfitting)

Polynomial fitting for $n \rightarrow m$: $\Rightarrow$ Fitting becomes interpolation: poor shape fidelity**Overfitting**

Fitting data with functions from a large space often produces poorer results (in terms of "shape preservation" and "generalization error") than the use of a smaller subspace for fitting.

§ 5.7.0.26 Data fitting with regularization

Idea: Suppress oscillations by "punishing"

large derivatives:

$$f \in \operatorname{argmin}_{g \in V} \left\{ \sum_{i=0}^n |g(t_i) - y_i|^2 + \alpha \int_a^b |g''(t)|^2 dt \right\}, \quad \alpha > 0 \quad (5.7.0.27)$$

$\nearrow$ $V \triangleq$ vector space of admissible functions
 $\searrow$ regularization parameter

 $V = \operatorname{Span} \{b_1, \dots, b_n\}$ [a basis]

$$\left[g = \sum_{j=1}^n (x)_j b_j \right] \Rightarrow g'' = \sum_{j=1}^n (x)_j b_j''$$

$$\phi(x) = \|Ax - y\|_2^2 + \alpha x^T M x \rightarrow \min, \quad \phi: \mathbb{R}^n \rightarrow \mathbb{R} \quad (5.7.0.28)$$

$$A := [b_j(t_i)]_{\substack{i=0, \dots, m \\ j=1, \dots, n}} \in \mathbb{R}^{m+1, n}, \quad M := \left[\int_a^b b_j''(t) b_k''(t) dt \right]_{j,k=1}^n \in \mathbb{R}^{n, n}. \quad (5.7.0.29)$$

$$\operatorname{grad} \phi(x) = 2A^T Ax - 2A^T y + 2\alpha Mx = 0,$$

 $\Leftrightarrow$

$$(A^T A + \alpha M)x = A^T y. \quad [\text{generalized normal equation}] \quad (5.7.0.30)$$

④

▷ Review questions 5.7.0.31

5

7