Modeling and Learning for Dynamical Systems

Lecture 6

Martin Enqvist

LINKÖPING UNIVERSITY

Static Linear Regression and Statistical Analysis

Linear Regression

The least-squares estimator of a linear regression model:

$$\hat{\theta}_{LS} = \arg\min_{\theta} \frac{1}{2}\|y - X\theta\|_2^2 = (X^T X)^{-1} X^T y$$

Assume true system $y = X\theta_0 + e$. Then:

- $E(\hat{\theta}_{LS}) = \theta_0$
- $\mathrm{Cov}(\hat{\theta}_{LS}) = (X^T X)^{-1} X^T \Sigma_e X (X^T X)^{-1}$
- Special case: $\Sigma_e = \sigma^2 I \Rightarrow \mathrm{Cov}(\hat{\theta}_{LS}) = \sigma^2 (X^T X)^{-1}$

Weighted least-squares estimator:

$$\hat{\theta}_{WLS} = (X^T \Sigma_e^{-1} X)^{-1} X^T \Sigma_e^{-1} y$$

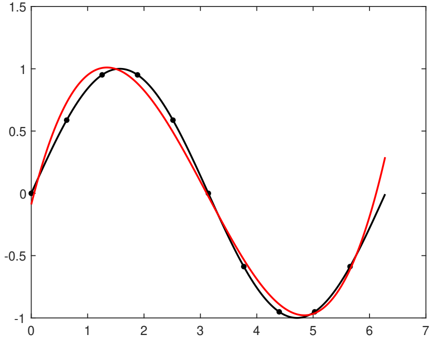
The LS/WLS estimator coincides with the maximum-likelihood (ML) estimator when e is normally distributed

Example: Sinusoid Approximated with Polynomial

$N = 10$, $\sigma = 0$ (no noise)

The fit is quite good inside the interval where we have data, which is reasonable since we have noisefree data and only approximation errors due to the polynomial not capturing the sinusoid exactly.

Estimate in red, true function in black solid, measurements as dots

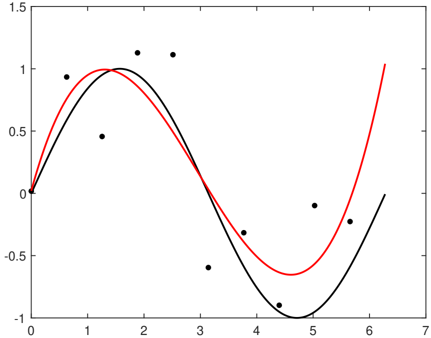


Example: Sinusoid Approximated with Polynomial...

$N = 10$, $\sigma = 0.5$

The fit is worse with noisy data and only a few measurements.

Estimate in red, true function in black solid, measurements as dots

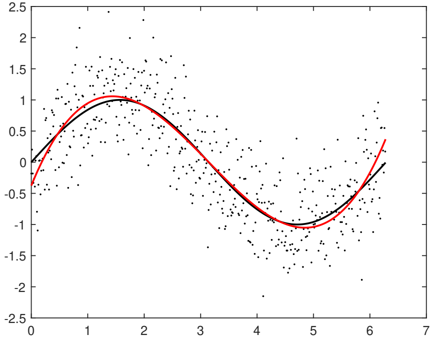


Example: Sinusoid Approximated with Polynomial...

$N = 500$, $\sigma = 0.5$

The fit becomes quite good with more measurements.

Estimate in red, true function in black solid, measurements as dots



Example 11.2 and 11.3 (Extended)

Consider a regression model

$$y_k = \varphi_k^T \theta_0 + e_k, \quad \theta_0 = \begin{pmatrix} 0.5 & 1.0 \end{pmatrix}^T$$

where $e_k \in \mathcal{N}(0,1)$ is white noise ($\Rightarrow$ LS = ML), and assume that $N = 25$ measurements are collected.

- LS estimate: $\hat{\theta}_{LS} = \begin{pmatrix} 0.3656 & 1.2079 \end{pmatrix}^T$ (Example 11.2)
- Regularized estimate with constraint $\theta_i \in [-1, 1]$: $\hat{\theta}_{R1.0} = \begin{pmatrix} 0.3661 & 1.0000 \end{pmatrix}^T$ (Example 11.3, clear effect on $\hat{\theta}_2$ since constraint is active)
- Regularized estimate with constraint $\theta_i \in [-1.5, 1.5]$: $\hat{\theta}_{R1.5} = \begin{pmatrix} 0.3656 & 1.2079 \end{pmatrix}^T$ (no effect since constraint is inactive)
- Regularized estimate using ridge regression, $\lambda = 1$: $\hat{\theta}_{RR} = \begin{pmatrix} 0.3363 & 1.1339 \end{pmatrix}^T$ (clear effect on both parameters)

Summary

Static Linear Regression and Statistical Analysis

- Linear regression models
- The least-squares estimator (expected value, covariance, weighted least-squares)
- Best linear unbiased estimator: Least-squares
- Regularization

www.liu.se

LINKÖPING
UNIVERSITY