Statistical and Computational Inverse Problems with Applications

Parts 3 & 4: Least squares estimation and regularization

Aku Seppänen

Inverse Problems Group
Department of Applied Physics
University of Eastern Finland
Kuopio, Finland

Jyväskylä Summer School
August 11-13, 2014

Linear LS problem

- Linear observation model

$$g = Kf + n$$

- LS-estimate

$$\hat{f}_{\mathrm{LS}} = \arg\min_{f} \left\{ \|g - Kf\|^2 \right\}$$

- The LS-estimate gets the form

$$\hat{f}_{\mathrm{LS}} = (K^{\mathrm{T}}K)^{-1}K^{\mathrm{T}}g$$

  if $(K^{\mathrm{T}}K)^{-1}$ exists, i.e. $\mathrm{rank}(K) = N$.
- See **Matlab examples 3.1 and 3.2**.

Minimum norm solution

- If $\mathrm{rank}(K) < N$, the LS-solution is not unique:
  - Let $\hat{f}$ be one solution of the LS problem. Then, for $\hat{f}$, $\|g - Kf\|$ is at minimum.
  - Let $w \in \mathrm{null}(K)$, $\|w\| = 1$, and let $a \in \mathbb{R}$. Then, $\|g - K(\hat{f} + aw)\| = \|g - K\hat{f} + a\underbrace{Kw}_{=0})\| = \|g - K\hat{f})\|$.
  - That is, $\hat{f} + aw$ is also a LS solution (for *any* $a \in \mathbb{R}$).
- The minimum norm estimate

$$\hat{f}_* = \arg\min_u \left\{ \|u\| \mid u = \arg\min_f \left\{ \|g - Kf\|^2 \right\} \right\}$$

- A matrix $K^\dagger$ for which

$$\hat{f}_* = K^\dagger g$$

  is called the *pseudoinverse* (**Matlab example 3.3**).
- However, the minimum norm solution does not necessarily have anything to do with the true f.

Generalized LS problem

- Assume again a linear observation model

$$g = Kf + n$$

with $\mathrm{rank}(K) = N$.

- Generalized LS-estimate

$$\hat{f}_{\mathrm{GLS}} = \arg\min_{f} \left\{ \|L(g - Kf)\|^2 \right\}$$

where L is a known (weighting) matrix.

- The estimate gets the form

$$\hat{f}_{\mathrm{GLS}} = (K^{\mathrm{T}} W K)^{-1} K^{\mathrm{T}} W g$$

where $W = L^{\mathrm{T}} L$.

Gauss-Markov estimate

- Assume again a linear observation model

$$g = Kf + n$$

with $\mathrm{rank}(K) = N$.
- It can be shown, than a linear estimate minimizing the criterion $\mathbb{E}\left\{\|f - \hat{f}\|^2\right\}$ is of the form

$$\hat{f}_{\mathrm{GM}} = \arg\min_f \left\{\|L(g - Kf)\|^2\right\}$$

where $L = \Gamma_n^{-1}$ (inverse of the noise covariance matrix).
- This corresponds to generalized LS with choice $W = \Gamma_n^{-1}$ and hence

$$\hat{f}_{\mathrm{GM}} = (K^{\mathrm{T}}\Gamma_n^{-1}K)^{-1}K^{\mathrm{T}}\Gamma_n^{-1}g$$

- See **Matlab examples 3.4 and 3.5**.

Non-linear LS problem

- Observation model

$$g = h(f) + n$$

- Generalized LS-estimate

$$\hat{f}_{\mathrm{GLS}} = \arg\min_{f} \left\{ \|L(g - h(f))\|^2 \right\}$$

- $\hat{f}_{\mathrm{GLS}}$ is solved iteratively. For example, Gauss-Newton method yields

$$\hat{f}_{k+1} = \hat{f}_k + a_k (J_k^{\mathrm{T}} W J_k)^{-1} (J_k^{\mathrm{T}} W (g - h(\hat{f}_k)))$$

  where J_k is the Jacobian matrix $J_k = \frac{\partial h}{\partial f}(\hat{f}_k)$ and a_k is a step length parameter.
- See **Matlab example 3.6.**

Generalized Tikhonov regularization

- Observation model

$$g = h(f) + n$$

- Assume that the problem is ill-posed, i.e.

$$\hat{f}_{\mathrm{GLS}} = \arg\min_f \left\{ \|L(g - h(f))\|^2 \right\}$$

is non-unique and/or unstable.

- Regularized solution

$$\hat{f}_\alpha = \arg\min_f \left\{ \|L(g - h(f))\|^2 + \alpha^2 \|L_\alpha(f - f_*)\|^2 \right\}$$

where α = regularization parameter and L_α = regularization matrix.

- Solution e.g. by Gauss-Newton method.

Generalized Tikhonov regularization

- In the case of a linear observation model $g = Kf + n$, the regularized solution

$$\hat{f}_\alpha = \arg\min_f \left\{ \|L(g - Kf)\|^2 + \alpha^2 \|L_\alpha(f - f_*)\|^2 \right\}$$

gets the form

$$\hat{f}_\alpha = (K^{\mathrm{T}} L^{\mathrm{T}} L K + \alpha L_\alpha^{\mathrm{T}} L_\alpha)^{-1} (K^{\mathrm{T}} L^{\mathrm{T}} L g + \alpha L_\alpha^{\mathrm{T}} L_\alpha f_*)$$

Generalized Tikhonov regularization

- The solution

$$\hat{f}_\alpha = \arg\min_f \left\{ \|g - Kf\|^2 + \alpha^2 \|f\|^2 \right\}$$

which corresponds to choices $L = I, L_\alpha = I,\ f_* = 0$ is referred to as (standard) Tikhonov regularization.

- A lot of research on selecting α (L-curve method, Morozov discrepancy principle,...)
- Other choices for the regularization matrix: e.g. a 1st or 2nd order difference matrix. See **Matlab example 4.1**: Regularized solution for numerical differentiation.