Lehrstuhl für Prozesstechnik
Dr. Ing. Adel Mhamdi

Lehrstuhl für Numerische Mathematik
Prof. Dr. A. Reusken

Model-based estimation methods, summer semester 2016

Sheet 6 (25th May, Thursday 18:15 - 19:45)

Problem 9 : SVD, Least squares

We reconsider Example 1.1.1 from the lecture notes. For $t \in [0,1]$ the following functions are given:

$$g(t) = \xi_1 t + \xi_2 e^t + \xi_3 t^3, \qquad \xi^* = (\xi_1, \xi_2, \xi_3)^T := (1.2,\ 0.6,\ 1.6)^T,$$
$$\hat{g}(t) = \xi_1 t + \xi_2 e^t + \xi_3 t^3 + \xi_4 \sin t, \qquad \hat{\xi}^* = (\xi_1, \xi_2, \xi_3, \xi_4)^T := (1.2,\ 0.6,\ 1.6,\ 0.9)^T.$$

Using discrete points $t_i = i\Delta t$, $i = 1, \ldots, 6$ for $\Delta t = 0.15$, we get (exact) data vectors $y, \hat{y} \in \mathbb{R}^6$ with $y_i = g(t_i)$ and $\hat{y}_i = \hat{g}(t_i)$, respectively. In the following we assume that the exact parameter values $\xi^*, \hat{\xi}^*$ are unknown and are determined by a least-square formulation using (measurement) data at t_i, $i = 1, \ldots, 6$.

- Construct matrices $A \in \mathbb{R}^{6\times 3}$ and $\hat{A} \in \mathbb{R}^{6\times 4}$ such that $A\xi^* = y$ and $\hat{A}\hat{\xi}^* = \hat{y}$ in MATLAB. Verify this by computing `A*xi - y` and `AHat*xiHat - yHat`.

- We choose perturbations $\delta y^1 := 10^{-2}\,(1, 0, -1, -1, -0.5, 1)^T$, $\delta y^2 := 10^{-2}\,(-1, 1, 1, -0.5, -2, 1)^T$, yielding perturbed data $y^i = y + \delta y^i$, $\hat{y}^i = \hat{y} + \delta y^i$, $i = 1, 2$, and consider the corresponding least-squares problems

$$\|A\xi - y^i\|_2 \to \min \qquad \text{and} \qquad \|\hat{A}\hat{\xi} - \hat{y}^i\|_2 \to \min, \qquad i = 1, 2.$$

  Solve these least-squares problems using the backslash operator, i.e., `xi = A\y`, and compute the errors $\|\xi^i - \xi^*\|_2$, $\|\hat{\xi}^i - \hat{\xi}^*\|_2$, $i = 1, 2$. Verify that you get the same errors as given in the lecture notes.

- To analyze the error propagation for the least-squares problems, compute the SVD of A and $\hat{A}$, respectively, by using the `svd` command in MATLAB. Discuss for both matrices: what is the largest possible error amplification, and for which kind of perturbation does it occur?

- Considering the least-squares problem for $\hat{g}$, represent the perturbations $\delta y^i \in \mathbb{R}^6$, $i = 1, 2$, in a suitable orthonormal basis. Why is there such a large error in $\hat{\xi}^2$, but only a mild error amplification in $\hat{\xi}^1$?

Problem 10 : Backwards heat equation

We consider the 1D (forwards) heat equation (cf. Section 1.3 in the lecture notes)

$$u_t(x,t) - \alpha\, u_{xx}(x,t) = 0 \qquad \text{for } x \in [0,L], t \in [0,T], \tag{1}$$

$$u(0,t) = U(L,t) = 0 \qquad \text{for } t \in [0,T], \tag{2}$$

$$u(x,0) = u_0(x) \qquad \text{for } x \in [0,L], \tag{3}$$

with length $L = \pi$, final time $T = 5$, heat conduction coefficient $\alpha = 10^{-3}$ and initial temperature $u_0(x) = \sin(x) + 0.2\sin(8x)$. For the backward heat equation, the initial condition (3) is replaced by a final condition $u(x,T) = u_T(x)$, $x \in [0,L]$.

Spatial discretization on a uniform 1D grid leads to the discrete problem

$$\vec{u}_t(t) - C\vec{u}(t) = 0 \qquad \text{for } t \in [0,T], \tag{4}$$

$$\vec{u}(0) = \vec{u}_0, \tag{5}$$

with

$$C = \frac{\alpha}{h^2}\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{n\times n}, \qquad \text{for mesh size } h = \frac{L}{n+1}.$$

- Compute the eigenvalues and eigenvectors of C for $n = 100$. Plot the eigenvectors corresponding to the smallest and largest eigenvalue, respectively.
- The solution of (4)–(5) is given by $\vec{u}(t) = \exp(-tC)\,\vec{u}_0$. Compute $\vec{u}_T = \vec{u}(T)$ with MATLAB. Adding a perturbation `1e-2*randn(n,1)` with normal distribution and variance $\sigma = 10^{-2}$ to $\vec{u}_T$ yields the measured final temperature $\vec{u}_T^{\text{meas}}$. Plot $\vec{u}_0, \vec{u}_T, \vec{u}_T^{\text{meas}}$ together in one figure (`hold on` ...`hold off`).
- Using $\vec{u}_T^{\text{meas}}$ as final condition, compute the estimated initial temperature $\vec{u}_0^{\text{est}} = \vec{u}(0)$ which is the solution of the corresponding backward heat equation. Plot $\vec{u}_0, \vec{u}_0^{\text{est}}$ together in one plot.
- Repeat the previous steps for $n = 50$ and $n = 25$. Which amount of ill-posedness do you observe on these coarser grids?

Problem 11 : TSVD, Tikhonov

We reconsider the least squares problem for $\hat{g}$ from Problem 9. For easier notation, we drop the '$\hat{\ }$' in the following.

$$g(t) = \xi_1 t + \xi_2 e^t + \xi_3 t^3 + \xi_4 \sin t, \qquad \xi^* = (\xi_1, \xi_2, \xi_3, \xi_4)^T := (1.2,\ 0.6,\ 1.6,\ 0.9)^T.$$

Using discrete points $t_i = i\Delta t$, $i = 1, \ldots, 6$ for $\Delta t = 0.15$, we get an (exact) data vector $y \in \mathbb{R}^6$ with $y_i = g(t_i)$. In the following we assume that the exact parameter values ξ^* are unknown and are determined by a least-square formulation using (measurement) data at t_i, $i = 1, \ldots, 6$.

We choose the perturbation $\delta y = \delta y^2 := 10^{-2}\,(-1, 1, 1, -0.5, -2, 1)^T$, yielding perturbed data $\tilde{y} = y + \delta y$ and consider the corresponding least-squares problem

$$\|A\tilde{\xi} - \tilde{y}\|_2 \to \min .$$

a.) Plot the exact parameters ξ^* and the estimated parameters $\tilde{\xi} = A^\dagger \tilde{y}$ together in one figure. Plot the exact data y, the perturbed data $\tilde{y}$ and the data fit $\bar{y} = A\tilde{\xi}$ together in one figure. What do you observe?

b.) Compute the SVD of A, $U^T A V = \Sigma$, and use the truncated SVD to get a regularized solution $\tilde{\xi}_{\text{TSVD}} = V \Sigma_{\text{reg}}^\dagger U^T \tilde{y}$, where in Σ_{reg} the smallest singular value σ_4 was replaced by zero. As before, plot the parameters $\xi^*, \tilde{\xi}_{\text{TSVD}}$ and the data $y, \tilde{y}, \bar{y} = A\tilde{\xi}_{\text{TSVD}}$ in respective figures. Discuss your results.

What happens if in Σ_{reg} also σ_3 is replaced by zero?

c.) Now we consider Tikhonov regularization, choosing $\alpha = 1$. Note that this is equivalent to replacing the normal equation $A^T A\, \xi = A^T y$ by the modified normal equation $(A^T A + \alpha^2 I)\xi_\alpha = A^T y$. As before, plot the parameters $\xi^*, \tilde{\xi}_\alpha$ and the data $y, \tilde{y}, \bar{y} = A\tilde{\xi}_\alpha$ in respective figures. Discuss your results.

d.) Obviously, the parameter choice $\alpha = 1$ was not optimal. To find a better choice, we try a parameter range $\alpha \in [10^{-3}, 10^3]$ (or `logspace(-3,3)` in MATLAB). Plot the parameter error $\|\tilde{\xi}_\alpha - \xi^*\|_2$ and the data error $\|A\tilde{\xi}_\alpha - \tilde{y}\|_2$ as functions of α together in a `loglog` plot. What would be a better parameter choice $\alpha = \alpha^*$? Repeat c.) for this improved parameter α^*.