TU Delft DCSC

Homework exercise I

Filtering and Identification (sc42025)

Hand in your solutions at the end of the lecture on **22-11-2017** (only hard-copy solutions are accepted).

Please highlight your final answer!

Exercise 1: Constrained least-squares problem

1. Consider the following unconstrained minimization problem:

$$\min_x \quad \|y - Fx\|_2^2$$

where $F \in \mathbb{R}^{m\times n}$ $(m > n)$ is a given full-rank matrix, and $y \in \mathbb{R}^m$ is a known vector. Find the optimal solution (denoted by x_u^*) to this unconstrained minimization problem.

2. Consider the following constrained minimization problem:

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & h(x) = 0 \end{aligned}$$

with $f(x) = \|y - Fx\|_2^2$ and the equality constraint $h(x) = Hx - g$, where $H \in \mathbb{R}^{l\times n}$ $(l \le n)$ is a given matrix and $g \in \mathbb{R}^l$ is a known vector. Introduce a new vector $\lambda \in \mathbb{R}^l$, called the Lagrange multiplier. This multiplier λ denotes the scaling factor between the parallel gradients of $f(x)$ and $h(x)$. The stationary point is given by the following equation:

$$\frac{\partial f(x)}{\partial x} + \frac{\partial h(x)}{\partial x}\lambda = 0$$

Resulting in the following linear system:

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} v \\ w \end{bmatrix}$$

Give an expression for the unknown matrices A, B, C and the unknown vectors v, w in terms of the given data F, y, H and g.

3. Assume $(C - B^T A^{-1} B)$ is invertible. Use the Schur complement to find the optimal solution (denoted by x_c^*) to this constrained minimization problem.

4. What does the solution x_c^* become for $\lambda = 0$?

Exercise 2: Least squares

In this first exercise, you are asked to derive two solutions for a least squares problem after which the statistic property between the solutions is compared. Given the set of equations:

$$\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} x + \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix}$$

denoted compactly as,

$$y = Fx + v.$$

1. Show that the following least squares solution:

$$\hat{x}_{\text{LS}} = \text{argmin}_x v^{\text{T}} v \ \text{ subject to } \ y = Fx + v$$

   is simply given by,

$$\hat{x}_{\text{LS}} = \frac{1}{N} \sum_{i=1}^{N} y(i)$$

2. When you denote the estimate $\hat{x}_{LS}$ obtained using N measurements by $\hat{x}_N$. And similarly denote the estimate obtained when using only the first $N-1$ measurements by $\hat{x}_{N-1}$. Then prove the following relationship between $\hat{x}_{N-1}$ and $\hat{x}_N$:

$$\hat{x}_N = \frac{N-1}{N} \hat{x}_{N-1} + \frac{1}{N} y(N). \tag{1}$$

3. Use the expression (1) to show that for increasing N, new measurements have a decreasing influence on the estimate $\hat{x}_N$.

4. Next assume that $v(i)$ for $i = 1 : N$ is zero-mean and has variance $E[v(i)^2] = 1$ for $i = 1 : N-1$ and $E[v(N)^2] = \sigma^2$. Then derive an estimate of the following weighted least squares problem:

$$\hat{x}_w = \text{argmin}_x \epsilon^T \epsilon \quad \text{subject to } y = Fx + \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \cdots & 0 & \sigma \end{bmatrix} \epsilon.$$

5. Prove that you can write $\hat{x}_w$ as:

$$\hat{x}_w = \frac{\sigma^2(N-1)}{\sigma^2(N-1)+1} \hat{x}_{N-1} + \frac{1}{\sigma^2(N-1)+1} y(N). \tag{2}$$

6. Use the expression (2) to show that when the N-th measurement is very accurate, the estimate $\hat{x}_w$ approximately equals that last very accurate measurement $y(N)$!