Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg December 18, 2017, 08:15-09:45, Freiburg

Surname: First Name: Matriculation number:

Subject: Programme: Bachelor ☐ Master ☐ Lehramt ☐ others ☐ Signature:

Please fill in your name above and tick exactly ONE box for the right answer of each question below.

1. Give an approximation of the covariance $\Sigma_{\hat{\theta}}$ of a maximum likelihood (ML) estimate. The model is given as $y_N = \Phi_N\theta + \epsilon_N$ with $\epsilon_N \sim \mathcal{N}(0, \Sigma_\epsilon)$, $Q_N = \Phi_N^\top \Phi_N$ and $L(\theta, y_N)$ is the negative log likelihood function. $\Sigma_{\hat{\theta}} \approx \ldots$

(a) ☐ $\Phi_N^+ \Sigma_{\epsilon_N} \Phi_N^{+\top}$	(b) ☐ $(\Phi_N^\top \Sigma_{\epsilon_N}^{-1} \Phi_N)^{-1}$	(c) ☐ Q_N^{-1}	(d) ☐ $\nabla_\theta^2 L(\theta, y_N)$

2. Given the probability density function $p_X(x) = \theta e^{-\theta x}$ for $x \geq 0$ (and 0 otherwise) with unknown θ and positive i.i.d. measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$ that are assumed to follow the above distribution, what is the minimisation problem you need to solve for a ML-estimate of θ? The problem is: $\min\limits_\theta \ldots$?

(a) ☐ $\|\theta e^{-\theta y(k)}\|_2^2$	(b) ☐ $-N\log(\theta) + \theta\sum_{k=1}^N y(k)$
(c) ☐ $\|y(k) - \theta e^{-\theta}\|_2^2$	(d) ☐ $-\log\sum_{k=1}^N \theta e^{-\theta y(k)}$

3. For the problem in the previous question, what is a lower bound on the covariance $\Sigma_{\hat{\theta}}$ for any unbiased estimator $\hat{\theta}(y_N)$, assuming that θ_0 is the true value? $\Sigma_{\hat{\theta}} \succeq \ldots$

(a) ☐ N/θ^2	(b) ☐ θ_0^2/N
(c) ☐ $\left(\int_{y_N} N\theta^{N-2}\exp[-\theta\sum_k y_k]\mathrm{d}y_N\right)^{-1}$	(d) ☐ $\int_{y_N} N\theta_0^{N-2}\exp[-\theta\sum_k y_k]\mathrm{d}y_N$

4. Suppose you are given the Fisher information matrix $M = \int_{y_N} \nabla_\theta^2 L(\theta_0, y_N) \cdot p(y_N|\theta_0)\mathrm{d}y_N$ of the corresponding problem, what is the relation with the covariance matrix $\Sigma_{\hat{\theta}}$ of your estimate $\hat{\theta}$?

5. Give the name of the theorem that provides us with the above result.

6. Given a set of measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$ with i.i.d. Gaussian noise and the linear model $y_N = \Phi\theta$, where $\Phi = [\varphi(1), \varphi(2), \ldots, \varphi(N)]^T$, which of the following minimisation problems is solved at each iteration step of the RLS algorithm to estimate the parameter $\hat{\theta}(N+1)$ after $N+1$ measurements? $\hat{\theta}(N+1) = \arg\min\limits_\theta \frac{1}{2}(\ldots)$

(a) ☐ $\|\theta - \hat{\theta}(N)\|_2^2 + \|y(N+1) - \varphi(N+1)^\top\theta\|_2^2$	(b) ☐ $\|y_N - \Phi_N \cdot \theta\|_{Q_N}^2$
(c) ☐ $\|\theta - \hat{\theta}(N)\|_{Q_N}^2 + \|y(N) - \varphi(N)^\top\theta\|_2^2$	(d) ☐ $\|y_{N+1} - \Phi_{N+1} \cdot \theta\|_2^2$

7. In L_1 estimation the measurement errors are assumed to follow a ... distribution and it is generally speaking more ... to outliers compared to L_2 estimation.

(a) ☐ Laplace, robust	(b) ☐ Gaussian, robust	(c) ☐ Gaussian, sensitive	(d) ☐ Laplace, sensitive

Points on page (max. 7) ☐

8. Write a general expression for an Auto Regressive model with Exogenous Inputs (ARX) with output errors:

9. The PDF of a random variable Y is given by $p(y) = \frac{1}{2\sqrt{2\pi}} \exp(-\frac{1}{2}\frac{\|y-\theta\|_2^2}{4})$, with unknown $\theta \in \mathbb{R}$. We obtained three measurements, $y(1) = 2$, $y(2) = 4$, and $y(3) = 6$. What is the minimizer θ^* of the negative log-likelihood function ?

(a) ☐ 6	(b) ☐ 3	(c) ☐ 5	(d) ☐ 4

10. Which of the following statements is **NOT** correct. Recursive Least Squares (RLS):

(a) ☐ implicitly assumes that there is only i.i.d. and Gaussian measurement noise	(b) ☐ computes an estimation with a computational cost independent of the number of past measurements
(c) ☐ can be used as an alternative to Maximum Likelihood Estimation	(d) ☐ can use prior knowledge on the estimated parameter θ

11. We want to estimate the resistance R of a new metal and we found in the only existing previous article that an estimate of R is given by $5[\Omega]$ with standard deviation 0.25Ω. Our own measurement apparatus sets a current I as a noise-free input, and measures the output voltage V which has Gaussian errors with a standard deviation of $0.1[V]$. Given a set of N measurements, $[V(1), \ldots, V(N)]$ obtained from a set $[I(1), \ldots, I(N)][A]$, what function is minimized by the Bayesian Maximum-A-Posteriori (MAP) estimator in this context? To simplify notation we assume that all variables are made unitless.

(a) ☐ $\frac{(R-10)^2}{0.5} + \sum_{i=1}^N \frac{(V(i)-I(i)R)^2}{0.1}$	(b) ☐ $\dfrac{(R-5)^2}{0.0625} + \sum_{i=1}^N 100(V(i) - I(i)R)^2$
(c) ☐ $0.5(R-10)^2 + \sum_{i=1}^N 0.1(V(i) - I(i)R)^2$	(d) ☐ $\frac{(R-10)^2}{0.5} + \sum_{i=1}^N \frac{(V(i)/I(i)-R)^2}{0.1}$

12. Which of the following models with input $u(k)$ and output $y(k)$ is **NOT** linear-in-the-parameters w.r.t. $\theta \in \mathbb{R}^2$?

(a) ☐ $y(k) = \theta_1\sqrt{u(k)} + \theta_2 u(k)$	(b) ☐ $y(k) = y(k-1) \cdot (\theta_1 + \theta_2 u(k))$
(c) ☐ $y(k) = \theta_1 u(k)^4 + \theta_2 \exp(u(k))$	(d) ☐ $y(k) = \theta_1 \exp(\theta_2 u(k))$

13. Which of the following models is time invariant?

(a) ☐ $\ddot{y}(t)^2 = u(t)^t + e^{u(t)}$	(b) ☐ $t \cdot \ddot{y}(t) = u(t)^3$	(c) ☐ $\dot{y}(t) = 5u(t) + t$	(d) ☐ $\dot{y}(t) = \sqrt{u(t)}$

14. By which of the following formulas is the joint distribution for N independent measurements $y_N \in \mathbb{R}^N$ given? $p(y_N|\theta) = \ldots$

(a) ☐ $\sum_{i=0}^N p(y(i)|\theta)$	(b) ☐ $\int_{y_N} p(y|\theta)\, dy$	(c) ☐ $\sum_{i=0}^N p(y(i)|\theta)^2$	(d) ☐ $\prod_{i=0}^N p(y(i)|\theta)$

15. Which of the following statements about Maximum A Posteriori (MAP) estimation is not true

(a) ☐ $\hat{\theta}_{\text{MAP}} = \arg\min_{\theta\in\mathbb{R}}[-\log(p(y_N|\theta)) - \log(p(\theta))]$	(b) ☐ MAP assumes a linear model
(c) ☐ MAP is a generalization of ML	(d) ☐ MAP requires a-priori knowledge on θ

16. Assume a model $h_i(\theta)$ and measurements y_i, $i = 1, \ldots, N$. The PDF to obtain a measurement y_i for a parameter θ is known to be proportional to $\exp(-|y_i - h_i(\theta)|)$ and measurement noises uncorrelated. What function of θ does the MLE minimize?

(a) ☐ $\sum_{i=1}^N |y_i - h_i(\theta)|$	(b) ☐ $|\sum_{i=1}^N y_i - \sum_{i=1}^N h_i(\theta)|$
(c) ☐ $\frac{1}{2}\|h(\theta) - y\|_2^2$	(d) ☐ $\sum_{i=1}^N (y_i - h_i(\theta))^2$

17. Regard the LLS estimate $\hat{\theta}$ minimizing $f(\theta) = \|y_N - \Phi_N\theta\|_2^2$, where measurements are generated by $y_N = \Phi_N\theta_0 + \epsilon_N$ with θ_0 the unknown true value, and $\epsilon_N = (\epsilon(1), \ldots, \epsilon(N))^\top$ the measurement errors (i.i.d., zero mean, variance σ^2, but not necessarily Gaussian). What would be the covariance matrix $\Sigma_{\hat{\theta}}$ of $\hat{\theta}$?

(a) ☐ $\sigma^2(\Phi_N^\top\Phi_N)^{-1}$	(b) ☐ $\sigma(\Phi_N^+)(\Phi_N^+)^\top$
(c) ☐ not computable	(d) ☐ $(\Phi_N^\top\sigma^2\Phi_N)^{-1}$

Points on page (max. 10)	