Modeling and System Identification – Microexam 1

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg November 28, 2016, 8:15-9:15, 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. What is the probability density function (PDF) $p_X(x)$ for a normally distributed random variable X with mean 3 and standard deviation 2? The answer is $p_X(x) = \frac{1}{\sqrt{2\pi 4}} \dots$

(a) ☐ $e^{-\frac{(x-3)^2}{8}}$	(b) ☐ $e^{-\frac{\lvert x-3 \rvert}{4}}$	(c) ☐ $e^{-\frac{(x+3)^2}{4}}$	(d) ☐ $e^{\frac{(x-3)^2}{8}}$

2. Which of the following statements does NOT hold for all PDFs $p(x)$?

(a) ☐ all others	(b) ☐ $p(x) \geq 0$	(c) ☐ $p(x) < 1$	(d) ☐ $\int_{-\infty}^{\infty} p(x) = 1$

3. What is the PDF of a random variable Y with uniform distribution on the interval $[0, \sqrt{2}]$? For $z \in [0, 1]$ it has the value:

(a) ☐ $p_Y(z) = \frac{1}{\sqrt{2}}$	(b) ☐ $p_Z(y) = \frac{1}{\sqrt{2}}$	(c) ☐ $p_Z(y) = \frac{\sqrt{2}}{y}$	(d) ☐ $p_Y(z) = 1$

4. What is the PDF normalization factor of an n-dimensional normally distributed variable Z with zero mean and covariance matrix $\Sigma \succ 0$? $p_Z(x) = \dots$

(a) ☐ $\frac{1}{\sqrt{(2\pi)^n}} \cdot \Sigma^{-1} \exp(-\frac{1}{2}x^\top \Sigma^{-1} x)$	(b) ☐ $\frac{1}{\sqrt{2\pi \det(\Sigma)}} \exp(-\frac{1}{2}x^\top \Sigma x)$
(c) ☐ $\frac{1}{\sqrt{2\pi \text{ trace}(\Sigma)}} \exp(-\frac{1}{2}x^\top \Sigma^\top x)$	(d) ☐ $\frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp(-\frac{1}{2}x^\top \Sigma^{-1} x)$

5. Regard a random variable $X \in \mathbb{R}^n$ with mean $\mu \in \mathbb{R}^n$ and covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$. For a fixed $b \in \mathbb{R}^n$ and $D, A \in \mathbb{R}^{m \times n}$, regard another random variable Y defined by $Y = AX + Db$. The mean of Y is given by $\mu_Y = \dots$

(a) ☐ $A\mu + Db$	(b) ☐ $A\Sigma^{-1}A^\top$	(c) ☐ $A^+ Db$	(d) ☐ $A\mu$

6. In question above, what is the covariance matrix of Y?

(a) ☐ $A^\top \Sigma A$	(b) ☐ $A \Sigma A^\top$	(c) ☐ $D\Sigma^{-1}D^\top$	(d) ☐ $D^+ \Sigma$

7. Regard a random variable $X \in \mathbb{R}^n$ with zero mean and covariance matrix Σ. Given a vector $c \in \mathbb{R}^n$, what is the mean of $Z = c^\top X X^\top c$?

(a) ☐ $\det(\Sigma)$	(b) ☐ $c^\top \text{trace}(\Sigma) c$	(c) ☐ $c^\top \Sigma c$	(d) ☐ $c^\top c \text{ trace}(\Sigma)$

8. Regard a random variable $\lambda \in \mathbb{R}$ with zero mean and standard deviation d. What is the mean of the random variable $Y = \lambda^2$?

(a) ☐ d^2	(b) ☐ $2\lambda d$	(c) ☐ 0	(d) ☐ d

9. Regard another scalar random variable that has variance $(d^2 - 2)$. What is its standard deviation?

(a) ☐ $\sqrt{d^2 - 2}$	(b) ☐ d	(c) ☐ 0	(d) ☐ $(d^2 - 2)^2$

10. Given a sequence of i.i.d. scalar random variables $X(1), \dots, X(N)$, each with mean μ and variance σ^2, what is the variance of the variable $Y = \frac{1}{N}\sum_{k=1}^{N} X(k)$? The answer is $\text{var}(Y) = \dots$

(a) ☐ $\frac{\sigma^2}{N}$	(b) ☐ $\frac{\sigma^2}{N-1}$	(c) ☐ $\frac{\sigma}{N}$	(d) ☐ $\frac{\sigma^2}{N^2}$

Points on page (max. 10)	

11. Which of the following functions is NOT convex on $x \in [-1, 1]$

(a) ☐ $\exp(-x)$	(b) ☐ $\sin^{-1}(x)$	(c) ☐ $x + 42$	(d) ☐ $-\cos(x)$

12. What is the minimizer x^* of the convex function $f : \mathbb{R} \to \mathbb{R}$, $\boxed{f(x) = 4 + \alpha x + \frac{1}{2}\beta x^2}$ with $\beta > 0$?

(a) ☐ $x^* = \frac{\beta}{4\alpha}$	(b) ☐ $x^* = -\frac{\alpha}{\beta}$	(c) ☐ $x^* = \frac{\beta}{\alpha}$	(d) ☐ $x^* = \frac{2\beta}{\alpha}$

13. For a matrix $\Phi \in \mathbb{R}^{N \times d}$ with rank d, what is its pseudo-inverse Φ^+ ?

(a) ☐ $(\Phi^\top \Phi)^{-1}\Phi^\top$	(b) ☐ $(\Phi^\top \Phi)^{-1}\Phi$	(c) ☐ $(\Phi\Phi^\top)^{-1}\Phi$	(d) ☐ $(\Phi\Phi^\top)^{-1}\Phi^\top$

14. What is the minimizer of the function $f : \mathbb{R}^n \to \mathbb{R}$, $\boxed{f(x) = \| - b + Dx\|_W^2}$ (with D of rank n and W positive semi-definite)? The answer is $x^* = \ldots$

(a) ☐ $(D^\top D)^{-1}D^\top b$	(b) ☐ $(DWD^\top)^{-1}DWb$
(c) ☐ $(DD^\top)^{-1}DWb$	(d) ☐ $(D^\top WD)^{-1}D^\top Wb$

15. Given a sequence of numbers $y(1), \ldots, y(N)$, what is the minimizer θ^* of the function $\boxed{f(\theta) = \sum_{k=1}^{N}(y(k) - 4\theta)^2}$? The answer is $\theta^* = \ldots$

(a) ☐ $\sum_{k=1}^{N} \frac{y(k)}{16N}$	(b) ☐ $\frac{\sum_{k=1}^{N} y(k)}{4N}$	(c) ☐ $\frac{1}{16N}\sum_{k=1}^{N} y(k)^2$	(d) ☐ $\frac{1}{4N}\sum_{k=1}^{N} y(k)^2$

16. Given a prediction model $\boxed{y(k) = \theta_1 + \frac{\theta_2}{2}x(k)^2 + \frac{\theta_3}{6}x(k)^3 + \epsilon(k)}$ with unknown parameter vector $\theta = (\theta_1, \theta_2, \theta_3)^\top$, and assuming i.i.d. noise $\epsilon(k)$ with zero mean, and given a sequence of N scalar input and output measurements $x(1), \ldots, x(N)$ and $y(1), \ldots, y(N)$, we want to compute the linear least squares (LLS) estimate $\hat{\theta}_N$ by minimizing the function $f(\theta) = \|y_N - \Phi_N\theta\|_2^2$. If $y_N = (y(1), \ldots, y(N))^\top$, how do we need to choose the matrix $\Phi_N \in \mathbb{R}^{N\times 3}$?

(a) ☐ $\begin{bmatrix} \frac{x(1)^2}{2} & 1 & \frac{x(1)^3}{6} \\ \vdots & \vdots & \vdots \\ \frac{x(N)^2}{2} & 1 & \frac{x(N)^3}{6} \end{bmatrix}$	(b) ☐ $\begin{bmatrix} 1 & 2x(1)^2 & 6x(1)^3 \\ \vdots & \vdots & \vdots \\ 1 & 2x(N)^2 & 6x(N)^3 \end{bmatrix}$	(c) ☐ $\begin{bmatrix} 1 & \frac{x(1)^2}{2} & \frac{x(1)^3}{6} \\ \vdots & \vdots & \vdots \\ 1 & \frac{x(N)^2}{2} & \frac{x(N)^3}{6} \end{bmatrix}$	(d) ☐ $\begin{bmatrix} 1 & x(1)^2 & x(1)^3 \\ \vdots & \vdots & \vdots \\ 1 & x(N)^2 & x(N)^3 \end{bmatrix}$

17. Which of the following formulas computes the covariance for a least squares estimator and a single experiment? $\hat{\Sigma}_{\hat{\theta}} = \ldots$

(a) ☐ $\frac{\|y_N - \Phi_N\hat{\theta}\|}{N-d}(\Phi_N\Phi_N^\top)^{-1}$	(b) ☐ $\frac{\|y_N - \Phi_N\hat{\theta}\|}{N-d}(\Phi_N^+\Phi_N)$
(c) ☐ $\frac{\|y_N - \Phi_N\hat{\theta}\|}{N-d}(\Phi_N^\top\Phi_N)^{-1}$	(d) ☐ $\Phi_N^+\sigma_{\epsilon_N}$

18. Given a set of measurements y_N following the model $y_N = \Phi_N\theta_0 + \epsilon_N$, where Φ_N is a regression matrix, θ_0 a vector with true parameter values and $\epsilon(k) \sim \mathcal{N}(0, \sigma_\epsilon^2)$ the noise contribution for $k = 1, ..., N$, we can compute the LLS estimator of the parameters θ as $\hat{\theta}_{\mathrm{LS}}$. Defining the covariance of $\hat{\theta}_{\mathrm{LS}}$ as $\Sigma_{\hat{\theta}}$, which of the following is NOT true?

(a) ☐ $\hat{\theta}_{\mathrm{LS}} = \Phi_N^+ y_N$	(b) ☐ $\Sigma_{\hat{\theta}} = \sigma_\epsilon^2(\Phi_N^+\Phi_N^{+\top})$
(c) ☐ $\hat{\theta}_{\mathrm{LS}} \sim \mathcal{N}(0, \Sigma_{\hat{\theta}})$	(d) ☐ $\hat{\theta}_{\mathrm{LS}}$ is a random variable

19. In the case given in the previous question, if the measurements y_N come from a single experiment, which condition on the noise do we need to require in order to be able to compute an estimate of σ_ϵ^2?

20. Imagine that the condition asked in the previous question is not met. We know that the noise has zero mean and covariance Σ_{ϵ_N}. What would be the covariance matrix $\Sigma_{\hat{\theta}}$ of the unweighted LLS estimate?

(a) ☐ $\Sigma_{\epsilon_N}(\Phi_N\Phi_N^\top)^{-1}$	(b) ☐ $\Phi_N^\top\Sigma_{\epsilon_N}^{-1}\Phi_N$	(c) ☐ $\Sigma_{\epsilon_N}^{-1}\Phi_N^\top\Phi_N$	(d) ☐ $\Phi_N^+\Sigma_{\epsilon_N}\Phi_N^{+\top}$

Points on page (max. 10)	