Modeling and System Identification
(Modellbildung und Systemidentifikation) – Exam

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg

March 20th, 2019, 14:00 - 17:00, Freiburg, Georges-Koehler-Allee 101 Room 026 & 036

Page	0	1	2	3	4	5	6	7	8	9	sum
Points on page (max)	4	7	9	11	8	5	9	7	4	0	64
Points obtained											
Intermediate sum											
Intermediate sum (max)	4	11	20	31	39	44	53	60	64	64	

Mark: Exam inspected on: Signature of examiner:

Surname: First name: Matriculation number:

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

Please fill in your name above. For the multiple choice questions, which give exactly one point, tick exactly one box for the right answer. For the text questions, give a short formula or text answer just below the question in the space provided, and, if necessary, write on the **backpage of the same sheet** where the question appears, and add a comment "see backpage". Do not add extra pages (for fast correction, all pages will be separated for parallelization). The exam is a closed book exam, i.e. no books or other material are allowed besides 2 sheet (with 4 pages) of notes and a non-programmable calculator. Some legal comments are found in a footnote.[1]

1. Give the probability density functions (PDF) $p_1(x)$ and $p_2(x)$ for normally distributed random variables with means $\mu_1 = -1$ cm and $\mu_2 = 0$ cm and standard deviations $\sigma_1 = 1$ cm and $\sigma_2 = 2$ cm. Add a sketch of *both in the same plot* including numbers and units on the x-axis:

   $p_1(x) =$ $p_2(x) =$

   Sketch:

4	

[1]WITHDRAWING FROM AN EXAMINATION: In case of illness, you must supply proof of your illness by submitting a medical report to the Examinations Office. Please note that the medical examination must be done at the latest on the same day of the missed exam. In case of illness while writing the exam please contact the supervisory staff, inform them about your illness and immediately see your doctor. The medical certificate must be submitted latest 3 days after the medical examination. More information's: http://www.tf.uni-freiburg.de/studies/exams/withdrawing_exam.html

CHEATING/DISTURBING IN EXAMINATIONS: A student who disrupts the orderly proceedings of an examination will be excluded from the remainder of the exam by the respective examiners or invigilators. In such a case, the written exam of the student in question will be graded as 'nicht bestanden' (5.0, fail) on the grounds of cheating. In severe cases, the Board of Examiners will exclude the student from further examinations.

points on page: 4

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

(a) ☐ $\int_{-\infty}^{\infty} p(x)\mathrm{d}x = 1$	(b) ☐ $p(x) < 1$
(c) ☐ $p(x) \geq 0$	(d) ☐ $\mathbb{E}(x) = \int_{-\infty}^{\infty} x \cdot p(x)\mathrm{d}x$

1	

3. Consider a multi-dimensional random variable $X \in \mathbb{R}^n$ with mean value $\mu \in \mathbb{R}^n$. What is its covariance? $\mathrm{cov}(X) = \ldots$

(a) ☐ $\mathbb{E}\{(X-\mu)\}^2$	(b) ☐ $\mathbb{E}\{(X-\mu)^2\}$
(c) ☐ $\mathbb{E}\{(X-\mu)(X-\mu)^\top\}$	(d) ☐ $\mathbb{E}\{(X-\mu)^\top(X-\mu)\}$

1	

4. 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 $c \in \mathbb{R}^n$, regard another random variable $Y \in \mathbb{R}^m$ defined by $Y = XX^\top + cc^\top$. The mean of Y is given by $\mu_Y = \ldots$

(a) ☐ $\Sigma + \mu\mu^\top + cc^\top$	(b) ☐ $\mu\mu^\top + cc^\top$	(c) ☐ $\Sigma + cc^\top$	(d) ☐ $\Sigma - \mu\mu^\top + cc^\top$

1	

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 $D, A \in \mathbb{R}^{m\times n}$ and a fixed $b \in \mathbb{R}^n$, regard another random variable $Z \in \mathbb{R}^m$ defined by $Z = AX + Db$. The covariance Z is given by $\Sigma_\mathrm{Z} = \ldots$

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

1	

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

(a) ☐ $x^2 + 4$	(b) ☐ $\sin(x) - x$
(c) ☐ $\exp(-x) + x^4$	(d) ☐ $-\cos(x)$

1	

7. You suspect your data to contain some outliers, thus you would use ... estimation which assumes your measurement errors follow a ... distribution.

(a) ☐ L_1, Laplace	(b) ☐ L_1, Gaussian	(c) ☐ L_2, Laplace	(d) ☐ L_2, Gaussian

1	

8. **Linear Least Squares**

Consider the linear model $y(k) = \theta_1 + \theta_2 x(k) + \varepsilon(k)$ and the vector of unknown parameters $\theta = [\theta_1, \theta_2]^\top$. The additive noise $\varepsilon(k)$ is assumed to be i.i.d. Gaussian and have zero mean. Given is a sequence of 3 scalar input and output measurements $x = [-1, 0, 1]^\top$ and $Y = [6, 6, 12]^\top$.

(a) State the least squares optimization problem that computes the minimizer θ^* as a sum of squared errors.

$$\theta^* = \arg\min_{\theta\in\mathbb{R}^2} \sum_{k=1}^{3}$$

1	

points on page: 7	

(b) The problem can be reformulated to be in the form: $\theta^* = \arg\min_{\theta \in \mathbb{R}^2} \frac{1}{2} \|Y - \mathbf{\Phi}\theta\|_2^2$.
State Y and $\mathbf{\Phi}$ explicitly and give their dimensions.

$Y =$ $\qquad\qquad$ $\mathbf{\Phi} =$

2	

(c) What is the analytical solution to the optimization problem in the previous question?

$\theta^* =$

1	

(d) Compute the optimizer θ^*. *Hint:* The inverse of a matrix $\mathbf{A} \in \mathbb{R}^{2\times 2}$ can be computed using the following formula:

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$\theta^* =$

3	

(e) State the formula that can be used to estimate the covariance $\hat{\Sigma}_{\theta^*}$.

$\hat{\Sigma}_{\theta^*} =$

1	

(f) Compute an estimate of the covariance $\hat{\Sigma}_{\theta^*}$.

$\hat{\Sigma}_{\theta^*} =$

2	

points on page: 9	

9. Given a LLS problem $\theta^* = \arg\min_{\theta \in \mathbb{R}^d} \|\mathbf{Y} - \mathbf{\Phi}\theta\|_2^2$, with $\mathbf{Y} \in \mathbb{R}^N, \theta \in \mathbb{R}^d, \mathbf{\Phi} \in \mathbb{R}^{N \times d}$ where $(\mathbf{\Phi}^\top \mathbf{\Phi}) \in \mathbb{R}^{d \times d}$ is non invertible.

   (a) How is such an estimation problem called? Why can that, in general, lead to a problem?

2	

   (b) State one technique which finds a unique solution in this case. Briefly Explain how it works and its drawbacks (2 sentences).

3	

10. **Weighted Least Squares**
    Consider the following WLS problem

$$\theta^* = \arg\min_{\theta} \|\mathbf{Y} - \mathbf{\Phi}\theta\|_{\mathbf{W}}^2$$

with $\mathbf{Y} \in \mathbb{R}^N$, $\theta \in \mathbb{R}^d$, $\mathbf{\Phi} \in \mathbb{R}^{N \times d}$, and $\mathbf{W} \in \mathbb{R}^{N \times N}$ is a symmetric positive definite matrix.

   (a) Which of the assumptions for LLS is removed in the WLS formulation?

1	

   (b) Motivate one case in which WLS is the preferred choice in comparison to LLS.

1	

   (c) Rewrite the WLS problem as an LLS problem by rewriting it in the form: $\arg\min_{\theta} \|\tilde{\mathbf{Y}} - \tilde{\mathbf{\Phi}}\theta\|_2^2$.
   Give an expression for $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{\Phi}}$.

   $\tilde{\mathbf{Y}} =$ $\tilde{\mathbf{\Phi}} =$

2	

   (d) Assume now that a different weight matrix is used $\tilde{\mathbf{W}} = 10\mathbf{W}$. How does this change the result of the optimizer θ^* and why?

2	

points on page: 11	

11. Let θ_{R} denote the *regularized* LLS estimator $\theta_{\mathrm{R}}(N)$ using L_2 regularization, i.e. $\theta_{\mathrm{R}}(N) = \arg\min_\theta \|Y_{\mathrm{N}} - \Phi_{\mathrm{N}}\theta\|_2^2 + \alpha\|\theta\|_2^2$, with $\alpha > 0$. Which of the following is **NOT** true?

(a) ☐ $\theta_{\mathrm{R}}(N)$ incorporates prior knowledge about θ.	(b) ☐ $\theta_{\mathrm{R}}(N)$ is biased.
(c) ☐ $\theta_{\mathrm{R}}(N)$ is asymptotically biased (for $N \to \infty$).	(d) ☐ $\theta_{\mathrm{R}}(N)$ can be computed analytically.

1

12. **Maximum Likelihood (ML)**
We assume that the annual return on an investment for a given stock can be modeled by Laplace distributions in the form:

$$p_X(x|\lambda) = \frac{1}{2\lambda} e^{-\dfrac{|x|}{\lambda}}, \qquad x \in \mathbb{R}, \lambda > 0,$$

with $\mathbb{E}\{|X|\} = \lambda$ being the expected value of return per year.

(a) Formulate the likelihood function $p(x_1, \ldots, x_N|\lambda)$ for N independent stock return observations $x_1, \ldots, x_N$.

$p(x_1, \ldots, x_N|\lambda) =$

1

(b) Compute the negative log-likelihood function $L(x_1, \ldots, x_N, \lambda)$ and simplify it.

$L(x_1, \ldots, x_N, \lambda) =$

2

(c) Show that the MLE for λ is given by: $\hat{\lambda}_{\mathrm{ML}} = \frac{1}{N}\sum_{i=1}^{N} |x_i|$
Hint: Use the first order necessary condition

2

(d) Is the ML estimator biased or unbiased? Prove your statement.

2

points on page: 8

13. Which of the following statements about Maximum A Posteriori (MAP) estimation is **NOT** true

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

1	

14. **Nonlinear Least Squares (NLS)**
NLS problems take the form: $\theta^* = \arg\min_\theta f(\theta)$, with objective function $f(\theta) = \frac{1}{2}\|R(\theta)\|_2^2$ and the residual function $R(\theta) : \mathbb{R}^d \to \mathbb{R}^N$.

(a) Since the objective's Hessian $B_{\text{exact}}(\theta) = \nabla_\theta^2 f(\theta)$ is generally expensive to compute and possibly indefinite it is often approximated. Suggest an approximation for the NLS problems including its definition, i.e., give a good approximation $B(\theta)$ of $B_{\text{exact}}(\theta)$ and state the its name.

$B(\theta) =$

2	

(b) Given the optimizer $\theta^* \in \mathbb{R}^d$ which was computed using a series of N measurements, give a formula to compute an estimate of the optimizer's covariance Σ_{θ^*}.

$\Sigma_{\theta^*} =$

2	

15. **Kalman Filter (KF) and Extended Kalman Filter (EKF)**
Let's consider the motion of a rocket in 2D space shortly after its launch. We would like the rocket to stay in vertical position, and we can control the thrusters that provide force orthogonal to the rocket's forward direction with the control u, in order to correct the rocket's orientation.

We model the state of the rocket by $x = \begin{bmatrix} p_1, & p_2, & \beta, & \omega \end{bmatrix}^\top \in \mathbb{R}^4$ where $p = [p_1,\ p_2]^\top$ denotes the 2-dimensional position of the rocket's center of gravity, βin radiant denotes the angle with respect to the vertical orientation and ω its angular velocity. The rocket's ODEs are given by

$$\dot{p}_1 = v\cos\beta \qquad \dot{p}_2 = v\sin\beta \qquad \dot{\beta} = \omega \qquad \dot{\omega} = \frac{uL}{2I}$$

where we assume the forward velocity v to be constant, L denotes the length of the rocket, and the rocket's moment of inertia is given by I.

We assume that we have perfect knowledge of the applied controls u. In addition, we have sensors which report the rocket's position p_1, p_2, orientation β, and angular velocity ω at any given moment in time.

Hint: If you don't find a solution to a question, continue to the remaining questions, using the general form.

points on page: 5	

(a) State the time derivative of the state vector using the given assumptions.

$\dot{x} =$

1	

(b) We want to track the rocket's state using the Kalman Filter. In preparation for this we need to linearize the model of the rocket. Please state a linearized model of the rocket of the form $\boxed{\dot{x} = Ax + Bu + d}$, with $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_u}$, $d \in \mathbb{R}^{n_x}$.
Hint: We assume that the controller keeps the rocket in almost vertical position, i.e. β is small, which allows us to make the following approximations $\sin\beta \approx \beta$ and $\cos\beta \approx 1$.

$A =$ $\qquad$ $B =$ $\qquad$ $d =$

3	

(c) Discretize the continuous time model found in the previous question using a single Euler step with step size h and specify a discrete time model of the form $\boxed{x_{k+1} = A_k x_k + B_k u_k + d_k}$, with $A_k \in \mathbb{R}^{n_x \times n_x}$, $B_k \in \mathbb{R}^{n_x \times n_u}$, $d_k \in \mathbb{R}^{n_x}$.

$A_k =$ $\qquad$ $B_k =$ $\qquad$ $d_k =$

3	

(d) We assume that both state and measurement model are subject to zero-mean additive Gaussian noise. Specify vectors and matrices $b_k \in \mathbb{R}^{n_x}$ and $C_k \in \mathbb{R}^{n_y \times n_x}$ such that

$$\begin{aligned} x_{k+1} &= A_k x_k + b_k + w_k \\ y_k &= C_k x_k + v_k \end{aligned}$$

where $w_k \sim \mathcal{N}(0, W_k)$ and $v_k \sim \mathcal{N}(0, V_k)$.

$b_k =$ $\qquad$ $C_k =$

2	

The Kalman Filter equations are given by

i. Prediction Step:

$$\begin{aligned} \hat{x}_{[k+1|k]} &= A_k \hat{x}_{[k|k]} + b_k \\ P_{[k+1|k]} &= A_k P_{[k|k]} A_k^\top + W_k \end{aligned}$$

ii. Innovation Update Step

$$\begin{aligned} P_{[k+1|k+1]} &= \left(P_{[k+1|k]}^{-1} + C_{k+1}^\top V_{k+1}^{-1} C_{k+1}\right)^{-1} \\ \hat{x}_{[k+1|k+1]} &= \hat{x}_{[k+1|k]} + P_{[k+1|k+1]} C_{k+1}^\top V_{k+1}^{-1} \left(y_{k+1} - C_k \hat{x}_{[k+1|k]}\right) \end{aligned}$$

points on page: 9	

(e) Briefly explain the role of W_k and V_k respectively and briefly describe how they should be chosen.

2	

(f) Which quantities do you need to initialize to start tracking the rocket's state of the rocket using the Kalman Filter (KF)? Do not choose real numbers here, but instead give an interpretation of these values.

2	

(g) Recall that we used a linear approximation of the $\cos$ and $\sin$ function in order to get a linear model. What is the advantage of using a linear model and the regular Kalman filter over using the more exact nonlinear model and the extended Kalman Filter (EKF)?

1	

(h) During operation of the rocket, your sensors suddenly stop working. What could you do to nevertheless track the state of the rocket until your sensors are operational again? What effect has this on the confidence ellipsoids?

2	

16. For each of the scenarios listed below, choose a state estimator that you would use and briefly justify your answer. Choose among the following answers:

- Linear Least Squares (LLS)
- Weighted linear least squares (WLS)
- Recursive least squares without forgetting factor (RLS)
- Recursive least squares with forgetting factor (RLSα)
- Kalman Filter (KF)
- Extended Kalman Filter (EKF)
- Moving Horizon Estimation (MHE)

points on page: 7	

(a) Linear state and measurement model with large Gaussian state and measurement noise, only little knowledge about the initial state of the system and very limited computational power available.

1	

(b) Nonlinear state and measurement model with large Gaussian state and measurement noise, only little knowledge about the initial state of the system.

1	

(c) Nonlinear state and measurement model with little Gaussian state and measurement noise, good knowledge of the initial state of the system and limited computational power available.

1	

(d) Linear state and nonlinear measurement model with Laplacian state and measurement noise, exact knowledge of the initial state.

1	

points on page: 4	

Empty page for calculations

points on page: 0	