Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg January 13, 2017, 10:15-11:45, 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. We want to assess the robustness of a particular smartphone brand hence, we would like to know the unknown probability θ that the phone breaks when it is dropped. We assume that the phone thrown onto the ground either breaks or has no damage. In an experiment we have dropped 100 smartphones and obtained 42 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the MLE estimate of θ ?

(a) ☒ $-42\log\theta - 58\log(1-\theta)$	(b) ☐ $-\log(42\theta) - \log(58(1-\theta))$
(c) ☐ $58\log\theta + 42\log(1-\theta)$	(d) ☐ $\log(58\theta) + \log(42(1-\theta))$

2. You are given a pendulum which is by nature a NON-LINEAR system and can be modeled by $y(t) = \theta_1\cos(\theta_2 t + \theta_3)$, where $y(t)$ are the measurements. Which of the following algorithms could you use to estimate the parameters θ?

(a) ☒ Maximum a Posteriori Estimation (MAP)	(b) ☐ Recursive Least Squares (RLS)
(c) ☐ Weighted Least Squares (WLS)	(d) ☐ Linear Least Squares (LLS)

3. Suppose now that the system given in the previous question can be approximated by a model that is linear in the parameters (LIP) . Which of the following algorithms could you use to estimate the parameters θ of this linear model without running into memory problems or high computational costs for a coninuous and infinite flow of measurement data?

(a) ☐ LLS	(b) ☐ MAP	(c) ☒ RLS	(d) ☐ WLS

4. You are asked to give a computationally efficient approximation of the covariance of the estimate computed in the previous question $\Sigma_{\hat{\theta}}$. 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. The covariance matrix can be approximated by $\Sigma_{\hat{\theta}} \approx \ldots$

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

5. Given the probability density function of the exponential distribution, $p_X(x) = \theta e^{-\theta x}$, with an unknown parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$, what is the right minimisation problem you need to solve for a maximum likelihood estimate of θ? The problem is: $\min_\theta \ldots$?

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

6. For the problem in the previous question, what is a lower bound on the covariance for any unbiased estimator $\hat{\theta}(y_N)$, assuming that θ_0 is the true value? The Fisher information matrix is defined as $M = \int_{y_N}\nabla_\theta^2 L(\theta_0, y_N)\cdot p(y_N|\theta_0)\mathrm{d}y_N$.

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

7. Suppose you are given the Fisher information matrix M of the corresponding problem, what is the relation with the covariance matrix $\Sigma_{\hat{\theta}}$ of your estimate $\hat{\theta}$?

$\Sigma_{\hat{\theta}} \succeq M^{-1}$

Points on page (max. 7) | |

8. Give the name of the theorem that provides us with the above result.
   Cramer-Rao Inequality

9. Given a set of measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$ from 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? $\hat{\theta}(N+1) = \arg\min_{\theta} \frac{1}{2}(\ldots)$

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

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

(a) [ ] Laplace, sensitive	(b) [ ] Laplace, robust	(c) [ ] Gaussian, robust	(d) [x] Gaussian, sensitive

11. 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, $y(1) = 3$, $y(2) = 6$, and $y(3) = 12$. What is the minimizer θ^* of the negative log-likelihood function ?

(a) [x] 7	(b) [ ] 6	(c) [ ] 9	(d) [ ] 4

12. Please give the ODE of a linear time invariant (LTI) system, with state vector x and input vector u. $\dot{x} = \ldots$
    $Ax + Bu$

13. Please identify the most general transfer function that still is a Auto Regressive Model with Exogenous Inputs (ARX) where $n = \max(n_a, n_b)$. $G(z) = \ldots$

(a) [ ] $\frac{b_0 z^n}{z^n + a_1 z^{n-1} + \ldots + a_n}$	(b) [ ] $\frac{z^n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(c) [x] $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(d) [ ] $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n}$

14. Identify most general transfer function that still is a Finite Impulse Response (FIR) model with $n = \max(n_a, n_b)$. $G(z) = \ldots$

(a) [ ] $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(b) [ ] $\frac{b_0 z^n}{z^n + a_1 z^{n-1} + \ldots + a_n}$	(c) [ ] $\frac{z^n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(d) [x] $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n}$

15. Which of the following model equations describes a FIR system with input u and output y? $y(k+1) = \ldots$

(a) [ ] $u(k+1) + y(k)$	(b) [ ] $u(k) \cdot y(k)$	(c) [ ] $u(k) + \sin(k \cdot \pi)$	(d) [x] $u(k) - 5 \cdot u(k-1)$

16. Which of the following dynamic models with inputs $u(t)$ and outputs $y(t)$ is **NEITHER** linear **NOR** affine.

(a) [ ] $\dot{y}(t) + \sin(t) = u(t)$	(b) [ ] $\dot{y}(t) = u(t) + t$	(c) [x] $\dot{y}(t) = \sqrt{t \cdot u(t)}$	(d) [ ] $t\dot{y}(t) = u(t) + 2$

17. 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) [ ] $y(k) = \theta_1 u(k)^2 + \theta_2 \exp(u(k))$	(b) [ ] $y(k) = \theta_1 \sqrt{u(k)} + \theta_2 u(k)$
(c) [ ] $y(k) = y(k-1) \cdot (\theta_1 + \theta_2 u(k))$	(d) [x] $y(k) = \theta_1 \exp(\theta_2 u(k))$

18. Which of the following models is time invariant?

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

19. 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) [ ] $\int_{y_N} p(y|\theta)\, dy$	(b) [ ] $\sum_{i=0}^{N} p(y(i)|\theta)^2$	(c) [x] $\prod_{i=0}^{N} p(y(i)|\theta)$	(d) [ ] $\sum_{i=0}^{N} p(y(i)|\theta)$

20. 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) [x] MAP assumes a linear model
(c) [ ] MAP is a generalization of ML	(d) [ ] MAP requires a-priori knowledge on θ

Points on page (max. 13)

Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg January 13, 2017, 10:15-11:45, 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. Given the probability density function of the exponential distribution, $p_X(x) = \theta e^{-\theta x}$, with an unknown parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$, what is the right minimisation problem you need to solve for maximum likelihood estimate of θ? The problem is: $\min_\theta \ldots$?

(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) ☐ $-\log\sum_{k=1}^N \theta e^{-\theta y(k)}$

2. For the problem in the previous question, what is a lower bound on the covariance for any unbiased estimator $\hat{\theta}(y_N)$, assuming that θ_0 is the true value? The Fisher information matrix is defined as $M = \int_{y_N} \nabla_\theta^2 L(\theta_0, y_N) \cdot p(y_N|\theta_0) \mathrm{d}y_N$.

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

3. You are given a pendulum which is by nature a NON-LINEAR system and can be modeled by $y(t) = \theta_1\cos(\theta_2 t + \theta_3)$, where $y(t)$ are the measurements. Which of the following algorithms could you use to estimate the parameters θ?

(a) ☒ Maximum a Posteriori Estimation (MAP)	(b) ☐ Linear Least Squares (LLS)
(c) ☐ Recursive Least Squares (RLS)	(d) ☐ Weighted Least Squares (WLS)

4. Suppose now that the system given in the previous question can be approximated by a model that is linear in the parameters (LIP) . Which of the following algorithms could you use to estimate the parameters θ of this linear model without running into memory problems or high computational costs for a coninuous and infinite flow of measurement data?

(a) ☐ MAP	(b) ☒ RLS	(c) ☐ LLS	(d) ☐ WLS

5. You are asked to give a computationally efficient approximation of the covariance of the estimate computed in the previous question $\Sigma_{\hat{\theta}}$. 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. The covariance matrix can be approximated by $\Sigma_{\hat{\theta}} \approx \ldots$

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

6. 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, $y(1) = 3$, $y(2) = 6$, and $y(3) = 12$. What is the minimizer θ^* of the negative log-likelihood function ?

(a) ☐ 4	(b) ☐ 6	(c) ☐ 9	(d) ☒ 7

7. Suppose you are given the Fisher information matrix M of the corresponding problem, what is the relation with the covariance matrix $\Sigma_{\hat{\theta}}$ of your estimate $\hat{\theta}$?
$\Sigma_{\hat{\theta}} \succeq M^{-1}$

8. Give the name of the theorem that provides us with the above result.
Cramer-Rao Inequality

Points on page (max. 8) | |

9. Which of the following models is time invariant?

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

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

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

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

12. Which of the following model equations describes a FIR system with input u and output y? $y(k+1) = \ldots$

(a) ☒ $u(k) - 5 \cdot u(k-1)$	(b) ☐ $u(k+1) + y(k)$	(c) ☐ $u(k) \cdot y(k)$	(d) ☐ $u(k) + \sin(k \cdot \pi)$

13. Which of the following dynamic models with inputs $u(t)$ and outputs $y(t)$ is **NEITHER** linear **NOR** affine.

(a) ☒ $\dot{y}(t) = \sqrt{t \cdot u(t)}$	(b) ☐ $\dot{y}(t) + \sin(t) = u(t)$	(c) ☐ $t\dot{y}(t) = u(t) + 2$	(d) ☐ $\dot{y}(t) = u(t) + t$

14. Given a set of measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$ from 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? $\hat{\theta}(N+1) = \arg\min_{\theta} \frac{1}{2}(\ldots)$

(a) ☒ $\|y_{N+1} - \Phi_{N+1} \cdot \theta\|_2^2$	(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) ☐ $\|\theta - \hat{\theta}(N)\|_2^2 + \|y(N+1) - \varphi(N+1)^\top \theta\|_2^2$

15. Please identify the most general transfer function that still is a Auto Regressive Model with Exogenous Inputs (ARX) where $n = \max(n_a, n_b)$. $G(z) = \ldots$

(a) ☐ $\frac{b_0 z^n}{z^n + a_1 z^{n-1} + \ldots + a_n}$	(b) ☐ $\frac{z^n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(c) ☒ $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(d) ☐ $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n}$

16. We want to assess the robustness of a particular smartphone brand hence, we would like to know the unknown probability θ that the phone breaks when it is dropped. We assume that the phone thrown onto the ground either breaks or has no damage. In an experiment we have dropped 100 smartphones and obtained 42 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the MLE estimate of θ ?

(a) ☐ $58 \log \theta + 42 \log(1 - \theta)$	(b) ☐ $-\log(42\theta) - \log(58(1 - \theta))$
(c) ☒ $-42 \log \theta - 58 \log(1 - \theta)$	(d) ☐ $\log(58\theta) + \log(42(1 - \theta))$

17. Please give the ODE of a linear time invariant (LTI) system, with state vector x and input vector u. $\dot{x} = \ldots$
$Ax + Bu$

18. 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) ☒ $y(k) = \theta_1 \exp(\theta_2 u(k))$	(b) ☐ $y(k) = y(k-1) \cdot (\theta_1 + \theta_2 u(k))$
(c) ☐ $y(k) = \theta_1 u(k)^2 + \theta_2 \exp(u(k))$	(d) ☐ $y(k) = \theta_1 \sqrt{u(k)} + \theta_2 u(k)$

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

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

20. Identify most general transfer function that still is a Finite Impulse Response (FIR) model with $n = \max(n_a, n_b)$. $G(z) = \ldots$

(a) ☐ $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(b) ☒ $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n}$	(c) ☐ $\frac{z^n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(d) ☐ $\frac{b_0 z^n}{z^n + a_1 z^{n-1} + \ldots + a_n}$

Points on page (max. 12) | |

Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg January 13, 2017, 10:15-11:45, 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. Suppose you are given the Fisher information matrix M of the corresponding problem, what is the relation with the covariance matrix $\Sigma_{\hat{\theta}}$ of your estimate $\hat{\theta}$?
   $\Sigma_{\hat{\theta}} \succeq M^{-1}$

2. Give the name of the theorem that provides us with the above result.
   Cramer-Rao Inequality

3. Please identify the most general transfer function that still is a Auto Regressive Model with Exogenous Inputs (ARX) where $n = \max(n_a, n_b)$. $G(z) = \dots$

(a) ☐ $\frac{z^n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(b) ☒ $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(c) ☐ $\frac{b_0 z^n}{z^n + a_1 z^{n-1} + \ldots + a_n}$	(d) ☐ $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n}$

4. Which of the following model equations describes a FIR system with input u and output y? $y(k+1) = \dots$

(a) ☒ $u(k) - 5 \cdot u(k-1)$	(b) ☐ $u(k) \cdot y(k)$	(c) ☐ $u(k) + \sin(k \cdot \pi)$	(d) ☐ $u(k+1) + y(k)$

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

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

6. Identify most general transfer function that still is a Finite Impulse Response (FIR) model with $n = \max(n_a, n_b)$. $G(z) = \dots$

(a) ☐ $\frac{z^n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(b) ☒ $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n}$	(c) ☐ $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(d) ☐ $\frac{b_0 z^n}{z^n + a_1 z^{n-1} + \ldots + a_n}$

7. We want to assess the robustness of a particular smartphone brand hence, we would like to know the unknown probability θ that the phone breaks when it is dropped. We assume that the phone thrown onto the ground either breaks or has no damage. In an experiment we have dropped 100 smartphones and obtained 42 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the MLE estimate of θ ?

(a) ☐ $\log(58\theta) + \log(42(1-\theta))$	(b) ☐ $58 \log \theta + 42 \log(1-\theta)$
(c) ☒ $-42 \log \theta - 58 \log(1-\theta)$	(d) ☐ $-\log(42\theta) - \log(58(1-\theta))$

8. Given a set of measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$ from 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? $\hat{\theta}(N+1) = \arg\min_{\theta} \frac{1}{2}(\ldots)$

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

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, $y(1) = 3$, $y(2) = 6$, and $y(3) = 12$. What is the minimizer θ^* of the negative log-likelihood function ?

(a) ☐ 9	(b) ☐ 6	(c) ☐ 4	(d) ☒ 7

Points on page (max. 9)

10. Please give the ODE of a linear time invariant (LTI) system, with state vector x and input vector u. $\dot{x} = \dots$
$Ax + Bu$

11. Given the probability density function of the exponential distribution, $p_X(x) = \theta e^{-\theta x}$, with an unknown parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \dots, y(N)]^T$, what is the right minimisation problem you need to solve for a maximum likelihood estimate of θ? The problem is: $\min\limits_{\theta} \dots$?

(a) [ ] $\Vert\theta e^{-\theta y(k)}\Vert_2^2$	(b) [x] $-N\log(\theta) + \theta\sum_{k=1}^{N} y(k)$
(c) [ ] $-\log\sum_{k=1}^{N} \theta e^{-\theta y(k)}$	(d) [ ] $\Vert y(k) - \theta e^{-\theta}\Vert_2^2$

12. For the problem in the previous question, what is a lower bound on the covariance for any unbiased estimator $\hat{\theta}(y_N)$, assuming that θ_0 is the true value? The Fisher information matrix is defined as $M = \int_{y_N} \nabla_\theta^2 L(\theta_0, y_N) \cdot p(y_N|\theta_0) \mathrm{d}y_N$.

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

13. Which of the following dynamic models with inputs $u(t)$ and outputs $y(t)$ is **NEITHER** linear **NOR** affine.

(a) [ ] $\dot{y}(t) + \sin(t) = u(t)$	(b) [ ] $\dot{y}(t) = u(t) + t$	(c) [x] $\dot{y}(t) = \sqrt{t \cdot u(t)}$	(d) [ ] $t\dot{y}(t) = u(t) + 2$

14. You are given a pendulum which is by nature a NON-LINEAR system and can be modeled by $y(t) = \theta_1\cos(\theta_2 t + \theta_3)$, where $y(t)$ are the measurements. Which of the following algorithms could you use to estimate the parameters θ?

(a) [x] Maximum a Posteriori Estimation (MAP)	(b) [ ] Weighted Least Squares (WLS)
(c) [ ] Linear Least Squares (LLS)	(d) [ ] Recursive Least Squares (RLS)

15. Suppose now that the system given in the previous question can be approximated by a model that is linear in the parameters (LIP) . Which of the following algorithms could you use to estimate the parameters θ of this linear model without running into memory problems or high computational costs for a coninuous and infinite flow of measurement data?

(a) [ ] LLS	(b) [ ] MAP	(c) [ ] WLS	(d) [x] RLS

16. You are asked to give a computationally efficient approximation of the covariance of the estimate computed in the previous question $\Sigma_{\hat{\theta}}$. 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. The covariance matrix can be approximated by $\Sigma_{\hat{\theta}} \approx \dots$

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

17. Which of the following models is time invariant?

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

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

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

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

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

20. 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) [ ] $y(k) = y(k-1) \cdot (\theta_1 + \theta_2 u(k))$	(b) [ ] $y(k) = \theta_1 u(k)^2 + \theta_2\exp(u(k))$
(c) [ ] $y(k) = \theta_1\sqrt{u(k)} + \theta_2 u(k)$	(d) [x] $y(k) = \theta_1\exp(\theta_2 u(k))$

Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg January 13, 2017, 10:15-11:45, 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. Which of the following models is time invariant?

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

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

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

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

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

4. You are given a pendulum which is by nature a NON-LINEAR system and can be modeled by $y(t) = \theta_1 \cos(\theta_2 t + \theta_3)$, where $y(t)$ are the measurements. Which of the following algorithms could you use to estimate the parameters θ?

(a) ☐ Weighted Least Squares (WLS)	(b) ☒ Maximum a Posteriori Estimation (MAP)
(c) ☐ Linear Least Squares (LLS)	(d) ☐ Recursive Least Squares (RLS)

5. Suppose now that the system given in the previous question can be approximated by a model that is linear in the parameters (LIP) . Which of the following algorithms could you use to estimate the parameters θ of this linear model without running into memory problems or high computational costs for a coninuous and infinite flow of measurement data?

(a) ☐ WLS	(b) ☐ MAP	(c) ☐ LLS	(d) ☒ RLS

6. You are asked to give a computationally efficient approximation of the covariance of the estimate computed in the previous question $\Sigma_{\hat{\theta}}$. 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. The covariance matrix can be approximated by $\Sigma_{\hat{\theta}} \approx \dots$

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

7. We want to assess the robustness of a particular smartphone brand hence, we would like to know the unknown probability θ that the phone breaks when it is dropped. We assume that the phone thrown onto the ground either breaks or has no damage. In an experiment we have dropped 100 smartphones and obtained 42 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the MLE estimate of θ ?

(a) ☐ $\log(58\theta) + \log(42(1-\theta))$	(b) ☒ $-42\log\theta - 58\log(1-\theta)$
(c) ☐ $-\log(42\theta) - \log(58(1-\theta))$	(d) ☐ $58\log\theta + 42\log(1-\theta)$

8. Given a set of measurements $y_N = [y(1), y(2), \dots, y(N)]^T$ from the linear model $y_N = \Phi\theta$, where $\Phi = [\varphi(1), \varphi(2), \dots, \varphi(N)]^T$, 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? $\hat{\theta}(N+1) = \arg\min_{\theta} \frac{1}{2}(\dots)$

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

Points on page (max. 8)

9. Which of the following dynamic models with inputs $u(t)$ and outputs $y(t)$ is **NEITHER** linear **NOR** affine.

(a) ☐ $\dot{y}(t)+\sin(t)=u(t)$	(b) ☐ $t\dot{y}(t) = u(t)+2$	(c) ☐ $\dot{y}(t)=u(t)+t$	(d) ☒ $\dot{y}(t)=\sqrt{t\cdot u(t)}$

10. 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) ☐ $y(k) = \theta_1 u(k)^2 + \theta_2 \exp(u(k))$	(b) ☐ $y(k) = y(k-1)\cdot(\theta_1 + \theta_2 u(k))$
(c) ☐ $y(k) = \theta_1\sqrt{u(k)} + \theta_2 u(k)$	(d) ☒ $y(k) = \theta_1 \exp(\theta_2 u(k))$

11. Which of the following model equations describes a FIR system with input u and output y? $y(k+1) = \ldots$

(a) ☒ $u(k) - 5\cdot u(k-1)$	(b) ☐ $u(k) + \sin(k\cdot\pi)$	(c) ☐ $u(k)\cdot y(k)$	(d) ☐ $u(k+1)+y(k)$

12. Suppose you are given the Fisher information matrix M of the corresponding problem, what is the relation with the covariance matrix $\Sigma_{\hat{\theta}}$ of your estimate $\hat{\theta}$?
$\Sigma_{\hat{\theta}} \succeq M^{-1}$

13. Give the name of the theorem that provides us with the above result.
Cramer-Rao Inequality

14. Identify most general transfer function that still is a Finite Impulse Response (FIR) model with $n = \max(n_a, n_b)$. $G(z) = \ldots$

(a) ☐ $\frac{b_0 z^n}{z^n + a_1 z^{n-1}+\ldots+a_n}$	(b) ☐ $\frac{b_0 z^n + b_1 z^{n-1}+\ldots+b_n}{a_0 z^n + a_1 z^{n-1}+\ldots+a_n}$	(c) ☐ $\frac{z^n}{a_0 z^n + a_1 z^{n-1}+\ldots+a_n}$	(d) ☒ $\frac{b_0 z^n + b_1 z^{n-1}+\ldots+b_n}{a_0 z^n}$

15. Please give the ODE of a linear time invariant (LTI) system, with state vector x and input vector u. $\dot{x} = \ldots$
$Ax + Bu$

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

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

17. Please identify the most general transfer function that still is a Auto Regressive Model with Exogenous Inputs (ARX) where $n = \max(n_a, n_b)$. $G(z) = \ldots$

(a) ☒ $\frac{b_0 z^n + b_1 z^{n-1}+\ldots+b_n}{a_0 z^n + a_1 z^{n-1}+\ldots+a_n}$	(b) ☐ $\frac{z^n}{a_0 z^n + a_1 z^{n-1}+\ldots+a_n}$	(c) ☐ $\frac{b_0 z^n}{z^n + a_1 z^{n-1}+\ldots+a_n}$	(d) ☐ $\frac{b_0 z^n + b_1 z^{n-1}+\ldots+b_n}{a_0 z^n}$

18. Given the probability density function of the exponential distribution, $p_X(x) = \theta e^{-\theta x}$, with an unknown parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$, what is the right minimisation problem you need to solve for a maximum likelihood estimate of θ? The problem is: $\min_{\theta} \ldots$?

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

19. For the problem in the previous question, what is a lower bound on the covariance for any unbiased estimator $\hat{\theta}(y_N)$, assuming that θ_0 is the true value? The Fisher information matrix is defined as $M = \int_{y_N} \nabla_\theta^2 L(\theta_0, y_N)\cdot p(y_N|\theta_0)\mathrm{d}y_N$.

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

20. 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, $y(1) = 3$, $y(2) = 6$, and $y(3) = 12$. What is the minimizer θ^* of the negative log-likelihood function ?

(a) ☐ 4	(b) ☒ 7	(c) ☐ 6	(d) ☐ 9

Points on page (max. 12)	

Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg January 13, 2017, 10:15-11:45, 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. Please give the ODE of a linear time invariant (LTI) system, with state vector x and input vector u. $\dot{x} = \dots$
   $Ax + Bu$

2. Please identify the most general transfer function that still is a Auto Regressive Model with Exogenous Inputs (ARX) where $n = \max(n_a, n_b)$. $G(z) = \dots$

(a) [ ] $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n}$	(b) [x] $\frac{b_0 z^n + b_1 z^{n-1} + \ldots + b_n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(c) [ ] $\frac{z^n}{a_0 z^n + a_1 z^{n-1} + \ldots + a_n}$	(d) [ ] $\frac{b_0 z^n}{z^n + a_1 z^{n-1} + \ldots + a_n}$

3. Suppose you are given the Fisher information matrix M of the corresponding problem, what is the relation with the covariance matrix $\Sigma_{\hat{\theta}}$ of your estimate $\hat{\theta}$?
   $\Sigma_{\hat{\theta}} \succeq M^{-1}$

4. Give the name of the theorem that provides us with the above result.
   Cramer-Rao Inequality

5. Which of the following models is time invariant?

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

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

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

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

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

8. Which of the following model equations describes a FIR system with input u and output y? $y(k+1) = \dots$

(a) [ ] $u(k) + \sin(k \cdot \pi)$	(b) [x] $u(k) - 5 \cdot u(k-1)$	(c) [ ] $u(k) \cdot y(k)$	(d) [ ] $u(k+1) + y(k)$

9. Which of the following dynamic models with inputs $u(t)$ and outputs $y(t)$ is **NEITHER** linear **NOR** affine.

(a) [x] $\dot{y}(t) = \sqrt{t \cdot u(t)}$	(b) [ ] $\dot{y}(t) + \sin(t) = u(t)$	(c) [ ] $\dot{y}(t) = u(t) + t$	(d) [ ] $t\dot{y}(t) = u(t) + 2$

10. We want to assess the robustness of a particular smartphone brand hence, we would like to know the unknown probability θ that the phone breaks when it is dropped. We assume that the phone thrown onto the ground either breaks or has no damage. In an experiment we have dropped 100 smartphones and obtained 42 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the MLE estimate of θ ?

(a) [ ] $58 \log\theta + 42 \log(1-\theta)$	(b) [x] $-42 \log\theta - 58 \log(1-\theta)$
(c) [ ] $-\log(42\theta) - \log(58(1-\theta))$	(d) [ ] $\log(58\theta) + \log(42(1-\theta))$

Points on page (max. 10)

11. Given the probability density function of the exponential distribution, $p_X(x) = \theta e^{-\theta x}$, with an unknown parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \dots, y(N)]^T$, what is the right minimisation problem you need to solve for a maximum likelihood estimate of θ? The problem is: $\min_{\theta} \dots$?

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

12. For the problem in the previous question, what is a lower bound on the covariance for any unbiased estimator $\hat{\theta}(y_N)$, assuming that θ_0 is the true value? The Fisher information matrix is defined as $M = \int_{y_N} \nabla_\theta^2 L(\theta_0, y_N) \cdot p(y_N|\theta_0)\mathrm{d}y_N$.

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

13. Given a set of measurements $y_N = [y(1), y(2), \dots, y(N)]^T$ from the linear model $y_N = \Phi\theta$, where $\Phi = [\varphi(1), \varphi(2), \dots, \varphi(N)]^T$, 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? $\hat{\theta}(N+1) = \arg\min_{\theta} \frac{1}{2}(\dots)$

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

14. Identify most general transfer function that still is a Finite Impulse Response (FIR) model with $n = \max(n_a, n_b)$. $G(z) = \dots$

(a) ☐ $\frac{b_0z^n + b_1z^{n-1} + \dots + b_n}{a_0z^n + a_1z^{n-1} + \dots + a_n}$	(b) ☒ $\frac{b_0z^n + b_1z^{n-1} + \dots + b_n}{a_0z^n}$	(c) ☐ $\frac{z^n}{a_0z^n + a_1z^{n-1} + \dots + a_n}$	(d) ☐ $\frac{b_0z^n}{z^n + a_1z^{n-1} + \dots + a_n}$

15. You are given a pendulum which is by nature a NON-LINEAR system and can be modeled by $y(t) = \theta_1\cos(\theta_2 t + \theta_3)$, where $y(t)$ are the measurements. Which of the following algorithms could you use to estimate the parameters θ?

(a) ☐ Recursive Least Squares (RLS)	(b) ☒ Maximum a Posteriori Estimation (MAP)
(c) ☐ Linear Least Squares (LLS)	(d) ☐ Weighted Least Squares (WLS)

16. Suppose now that the system given in the previous question can be approximated by a model that is linear in the parameters (LIP) . Which of the following algorithms could you use to estimate the parameters θ of this linear model without running into memory problems or high computational costs for a coninuous and infinite flow of measurement data?

(a) ☐ WLS	(b) ☐ MAP	(c) ☒ RLS	(d) ☐ LLS

17. You are asked to give a computationally efficient approximation of the covariance of the estimate computed in the previous question $\Sigma_{\hat{\theta}}$. 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. The covariance matrix can be approximated by $\Sigma_{\hat{\theta}} \approx \dots$

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

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

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

19. 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) ☒ $y(k) = \theta_1\exp(\theta_2 u(k))$	(d) ☐ $y(k) = \theta_1 u(k)^2 + \theta_2\exp(u(k))$

20. 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, $y(1) = 3$, $y(2) = 6$, and $y(3) = 12$. What is the minimizer θ^* of the negative log-likelihood function ?

(a) ☐ 9	(b) ☐ 4	(c) ☐ 6	(d) ☒ 7

Points on page (max. 10)	