Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg December 19, 2018, 14:10-14:55, 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 would like to know the unknown probability θ that a phone does **NOT** break 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 23 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the maximum likelihood (ML) estimate of θ ?

(a) ☐ $-\log(77\theta) - \log(23(1-\theta))$	(b) ☐ $23\log\theta + 77\log(1-\theta)$
(c) ☒ $-77\log\theta - 23\log(1-\theta)$	(d) ☐ $\log(23\theta) + \log(77(1-\theta))$

2. You are given a pendulum which is by nature a nonlinear 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 should 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)

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 continuous and infinite flow of measurement data?

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

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

5. Given the probability density function $p_X(x) = \theta x \exp(-\theta x)$, with parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$, which minimisation problem you need to solve to get a ML-estimate of θ? The problem is: $\min_\theta \ldots$?

(a) ☐ $\sum_k \|\theta y(k) \exp(-\theta y(k))\|_2^2$	(b) ☐ $N\log(\theta) + \theta \sum_k y(k)$
(c) ☐ $-\log\left(\sum_k \|\theta y(k) \exp(-\theta y(k))\|_2^2\right)$	(d) ☒ $-N\log(\theta) + \theta \sum_k y(k)$

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_0^2	(b) ☐ $\int_{y_N} N\theta_0^{N-2} \exp[-\theta_0 \sum_k y_k] \mathrm{d}y_N$
(c) ☒ θ_0^2/N	(d) ☐ $\int_{y_N} N\theta_0^{N-2} \left(\sum_k y(k)\right) \exp[-\theta_0 \sum_k y_k] \mathrm{d}y_N$

7. Let θ_R denote the *regularized* LLS estimator using L_2 regularization. Which of the following is **NOT** true?

(a) ☐ θ_R incorporates prior knowledge about θ.	(b) ☐ θ_R can be computed analytically.
(c) ☒ θ_R is asymptotically biased.	(d) ☐ θ_R is biased.

8. We use the Gauss-Newton (GN) algorithm to solve a nonlinear estimation problem. Which of the following statements is **NOT** true *in general*?

(a) ☐ The inverse of the GN Hessian approximates Σ_θ.	(b) ☐ The idea of GN is to linearize the residual function.
(c) ☐ GN uses a Hessian approximation.	(d) ☒ GN finds the global minimizer of the objective function.

Points on page (max. 8)	

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 and c are correct!]

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

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

(a) [ ] L_1, Gaussian	(b) [x] L_1, Laplace	(c) [ ] L_2, Laplace	(d) [ ] L_2, Gaussian

11. Assume you solved a nonlinear parameter estimation problem using the Gauss-Newton algorithm. How can you check that your assumptions on the noise distribution are correct?
Plot the entries of the residual vector, $R(\theta^*)_i$, $i = 1, \ldots, N$, as a histogram and check if it looks like a zero-mean Gaussian.

12. For a discrete time LTI system $x_{k+1} = Ax_k + Bu_k$, $k = 0, 1$, explicitly state the forward simulation map $f_{\text{sim}} : \mathbb{R}^{n_x + 2n_u} \to \mathbb{R}^{3n_x}$, $(x_0, u_0, u_1) \mapsto (x_0, x_1, x_2)$.

$$f(x_0, u_0, u_1) = \begin{bmatrix} x_0 \\ Ax_0 + Bu_0 \\ A^2x_0 + ABu_0 + Bu_1 \end{bmatrix}$$

13. Please identify the most general system equation that still is a Auto Regressive Model with Exogenous Inputs (ARX).

(a) [ ] $y(k) = -a_1 y(k-1) - \ldots - a_{n_a} y(k - n_a)$	(b) [ ] $y(k) = b_0 u(k) + \ldots + b_{n_b} u(k - n_b)$
(c) [ ] $a_0 y(k) + a_1 y^2(k-1) + \ldots + a_{n_a} y^{n_a+1}(k - n_a) = b_0 u(k) + \ldots + b_0 u(k - n_b)$	(d) [x] $a_0 y(k) + a_1 y(k-1) + \ldots + a_{n_a} y(k - n_a) = b_0 u(k) + \ldots + b_{n_b} u(k - n_b)$

14. Which one of the following statements is **NOT** true for FIR models:

(a) [x] The impulse response is constant.	(b) [ ] The output does not depend on previous outputs.
(c) [ ] Output error minimization is a convex problem.	(d) [ ] They are a special class of ARX models

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

(a) [ ] $u(k) + \sin(k \cdot \pi)^2$	(b) [ ] $u(k) \cdot y(k)$	(c) [x] $u(k) - \sqrt{\pi} u(k-2)$	(d) [ ] $u(k+1) + y(k)$

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

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

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

18. Which of the following models is time invariant?

(a) [ ] $\ddot{y}(t)^2 = u(t)^t + e^{u(t)}$	(b) [x] $\dot{y}(t) = \sqrt{u(t)} + 1$	(c) [ ] $\dot{y}(t) = -3u(t) + t^2$	(d) [ ] $t \cdot \ddot{y}(t) = u(t)^3$

19. With which of the following formulas you can **NOT** compute the conditional joint distribution of N independent measurements $y_N = (y_1, \ldots, y_N)$ given θ? $p(y_N|\theta) \neq \ldots$

(a) [ ] $\int p(y_N|x_N, \theta) p(x_N) \mathrm{d}x_N$	(b) [ ] $\prod_i p(y(i)|\theta)$
(c) [x] $\int p(y_N|\theta) p(\theta) \mathrm{d}\theta$	(d) [ ] $\exp\left(\sum_{i=0}^{N} \ln(p(y(i)|\theta))\right)$

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

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

Points on page (max. 12)	

Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg December 19, 2018, 14:10-14:55, 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 $p_X(x) = \theta x \exp(-\theta x)$, with parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$, which minimisation problem you need to solve to get a ML-estimate of θ? The problem is: $\min_\theta \ldots$?

(a) ☒ $-N\log(\theta) + \theta \sum_k y(k)$	(b) ☐ $\sum_k \|\theta y(k)\exp(-\theta y(k))\|_2^2$
(c) ☐ $N\log(\theta) + \theta \sum_k y(k)$	(d) ☐ $-\log\left(\sum_k \|\theta y(k)\exp(-\theta y(k))\|_2^2\right)$

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

3. You are given a pendulum which is by nature a nonlinear 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 should you use to estimate the parameters θ?

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

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 continuous and infinite flow of measurement data?

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

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

6. Assume you solved a nonlinear parameter estimation problem using the Gauss-Newton algorithm. How can you check that your assumptions on the noise distribution are correct?
Plot the entries of the residual vector, $R(\theta^*)_i$, $i = 1, \ldots, N$, as a histogram and check if it looks like a zero-mean Gaussian.

7. Let θ_R denote the *regularized* LLS estimator using L_2 regularization. Which of the following is **NOT** true?

(a) ☐ θ_R is biased.	(b) ☒ θ_R is asymptotically biased.
(c) ☐ θ_R can be computed analytically.	(d) ☐ θ_R incorporates prior knowledge about θ.

8. We use the Gauss-Newton (GN) algorithm to solve a nonlinear estimation problem. Which of the following statements is **NOT** true *in general*?

(a) ☐ The idea of GN is to linearize the residual function.	(b) ☒ GN finds the global minimizer of the objective function.
(c) ☐ GN uses a Hessian approximation.	(d) ☐ The inverse of the GN Hessian approximates Σ_θ.

Points on page (max. 8) ☐

9. Which of the following models is time invariant?

(a) [x] $\dot{y}(t) = \sqrt{u(t)} + 1$	(b) [ ] $\dot{y}(t) = -3u(t) + t^2$	(c) [ ] $\ddot{y}(t)^2 = u(t)^t + e^{u(t)}$	(d) [ ] $t \cdot \ddot{y}(t) = u(t)^3$

10. With which of the following formulas you can **NOT** compute the conditional joint distribution of N independent measurements $y_N = (y_1, \ldots, y_N)$ given θ? $p(y_N|\theta) \neq \ldots$

(a) [ ] $\int p(y_N|x_N, \theta)p(x_N)\mathrm{d}x_N$	(b) [ ] $\prod_i p(y(i)|\theta)$
(c) [x] $\int p(y_N|\theta)p(\theta)\mathrm{d}\theta$	(d) [ ] $\exp\left(\sum_{i=0}^{N} \ln(p(y(i)|\theta))\right)$

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

(a) [ ] The MAP estimator is biased.	(b) [ ] MAP is a generalization of ML.
(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.

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

(a) [ ] $u(k) + \sin(k \cdot \pi)^2$	(b) [ ] $u(k) \cdot y(k)$	(c) [ ] $u(k+1) + y(k)$	(d) [x] $u(k) - \sqrt{\pi}u(k-2)$

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

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

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 and c are correct!]

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

15. Please identify the most general system equation that still is a Auto Regressive Model with Exogenous Inputs (ARX).

(a) [ ] $y(k) = -a_1 y(k-1) - \ldots - a_{n_a} y(k-n_a)$	(b) [ ] $y(k) = b_0 u(k) + \ldots + b_{n_b} u(k-n_b)$
(c) [x] $a_0 y(k) + a_1 y(k-1) + \ldots + a_{n_a} y(k-n_a) = b_0 u(k) + \ldots + b_{n_b} u(k-n_b)$	(d) [ ] $a_0 y(k) + a_1 y^2(k-1) + \ldots + a_{n_a} y^{n_a+1}(k-n_a) = b_0 u(k) + \ldots + b_0 u(k-n_b)$

16. We would like to know the unknown probability θ that a phone does **NOT** break 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 23 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the maximum likelihood (ML) estimate of θ ?

(a) [ ] $23\log\theta + 77\log(1-\theta)$	(b) [x] $-77\log\theta - 23\log(1-\theta)$
(c) [ ] $\log(23\theta) + \log(77(1-\theta))$	(d) [ ] $-\log(77\theta) - \log(23(1-\theta))$

17. For a discrete time LTI system $x_{k+1} = Ax_k + Bu_k$, $k = 0, 1$, explicitly state the forward simulation map $f_{\text{sim}} : \mathbb{R}^{n_x + 2n_u} \to \mathbb{R}^{3n_x}$, $(x_0, u_0, u_1) \mapsto (x_0, x_1, x_2)$.

$$f(x_0, u_0, u_1) = \begin{bmatrix} x_0 \\ Ax_0 + Bu_0 \\ A^2x_0 + ABu_0 + Bu_1 \end{bmatrix}$$

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

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

(a) [x] L_1, Laplace	(b) [ ] L_1, Gaussian	(c) [ ] L_2, Gaussian	(d) [ ] L_2, Laplace

20. Which one of the following statements is **NOT** true for FIR models:

(a) [ ] They are a special class of ARX models	(b) [x] The impulse response is constant.
(c) [ ] Output error minimization is a convex problem.	(d) [ ] The output does not depend on previous outputs.

Points on page (max. 12)	

Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg December 19, 2018, 14:10-14:55, 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. Let θ_R denote the *regularized* LLS estimator using L_2 regularization. Which of the following is **NOT** true?

(a) [ ] θ_R is biased.	(b) [ ] θ_R incorporates prior knowledge about θ.
(c) [ ] θ_R can be computed analytically.	(d) [x] θ_R is asymptotically biased.

2. We use the Gauss-Newton (GN) algorithm to solve a nonlinear estimation problem. Which of the following statements is **NOT** true *in general*?

(a) [ ] GN uses a Hessian approximation.	(b) [ ] The idea of GN is to linearize the residual function.
(c) [x] GN finds the global minimizer of the objective function.	(d) [ ] The inverse of the GN Hessian approximates Σ_θ.

3. Please identify the most general system equation that still is a Auto Regressive Model with Exogenous Inputs (ARX).

(a) [ ] $y(k) = b_0 u(k) + \ldots + b_{n_b} u(k - n_b)$	(b) [ ] $y(k) = -a_1 y(k-1) - \ldots - a_{n_a} y(k - n_a)$
(c) [x] $a_0 y(k) + a_1 y(k-1) + \ldots + a_{n_a} y(k-n_a) = b_0 u(k) + \ldots + b_{n_b} u(k-n_b)$	(d) [ ] $a_0 y(k) + a_1 y^2(k-1) + \ldots + a_{n_a} y^{n_a+1}(k-n_a) = b_0 u(k) + \ldots + b_0 u(k-n_b)$

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

(a) [x] $u(k) - \sqrt{\pi} u(k-2)$	(b) [ ] $u(k) \cdot y(k)$	(c) [ ] $u(k) + \sin(k \cdot \pi)^2$	(d) [ ] $u(k+1) + y(k)$

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

(a) [ ] L_1, Gaussian	(b) [ ] L_2, Laplace	(c) [x] L_1, Laplace	(d) [ ] L_2, Gaussian

6. Which one of the following statements is **NOT** true for FIR models:

(a) [ ] Output error minimization is a convex problem.	(b) [ ] They are a special class of ARX models
(c) [ ] The output does not depend on previous outputs.	(d) [x] The impulse response is constant.

7. We would like to know the unknown probability θ that a phone does **NOT** break 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 23 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the maximum likelihood (ML) estimate of θ ?

(a) [x] $-77 \log \theta - 23 \log(1-\theta)$	(b) [ ] $\log(23\theta) + \log(77(1-\theta))$
(c) [ ] $-\log(77\theta) - \log(23(1-\theta))$	(d) [ ] $23 \log \theta + 77 \log(1-\theta)$

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 and c are correct!]

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

9. Assume you solved a nonlinear parameter estimation problem using the Gauss-Newton algorithm. How can you check that your assumptions on the noise distribution are correct?
Plot the entries of the residual vector, $R(\theta^*)_i$, $i = 1, \ldots, N$, as a histogram and check if it looks like a zero-mean Gaussian.

10. For a discrete time LTI system $x_{k+1} = Ax_k + Bu_k$, $k = 0, 1$, explicitly state the forward simulation map
$f_{\text{sim}} : \mathbb{R}^{n_x + 2n_u} \to \mathbb{R}^{3n_x}$, $(x_0, u_0, u_1) \mapsto (x_0, x_1, x_2)$.

Points on page (max. 9)

$$f(x_0, u_0, u_1) = \begin{bmatrix} x_0 \\ Ax_0 + Bu_0 \\ A^2x_0 + ABu_0 + Bu_1 \end{bmatrix}$$

11. Given the probability density function $p_X(x) = \theta x \exp(-\theta x)$, with parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \dots, y(N)]^T$, which minimisation problem you need to solve to get a ML-estimate of θ? The problem is: $\min_\theta \dots$?

(a) ☐ $N \log(\theta) + \theta \sum_k y(k)$	(b) ☐ $-\log\left(\sum_k \|\theta y(k) \exp(-\theta y(k))\|_2^2\right)$
(c) ☐ $\sum_k \|\theta y(k) \exp(-\theta y(k))\|_2^2$	(d) ☒ $-N \log(\theta) + \theta \sum_k y(k)$

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

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

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

14. You are given a pendulum which is by nature a nonlinear 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 should you use to estimate the parameters θ?

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

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 continuous and infinite flow of measurement data?

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

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

17. Which of the following models is time invariant?

(a) ☒ $\dot{y}(t) = \sqrt{u(t)} + 1$	(b) ☐ $\ddot{y}(t)^2 = u(t)^t + e^{u(t)}$	(c) ☐ $t \cdot \ddot{y}(t) = u(t)^3$	(d) ☐ $\dot{y}(t) = -3u(t) + t^2$

18. With which of the following formulas you can **NOT** compute the conditional joint distribution of N independent measurements $y_N = (y_1, \dots, y_N)$ given θ? $p(y_N|\theta) \neq \dots$

(a) ☐ $\int p(y_N|x_N, \theta) p(x_N) \mathrm{d}x_N$	(b) ☐ $\prod_i p(y(i)|\theta)$
(c) ☐ $\exp\left(\sum_{i=0}^N \ln(p(y(i)|\theta))\right)$	(d) ☒ $\int p(y_N|\theta) p(\theta) \mathrm{d}\theta$

19. 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) ☐ The MAP estimator is biased.	(d) ☐ $\hat{\theta}_{\text{MAP}} = \arg\min_{\theta \in \mathbb{R}}[-\log(p(y_N|\theta)) - \log(p(\theta))]$

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

Points on page (max. 11) | |

Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg December 19, 2018, 14:10-14:55, 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) [ ] $t \cdot \ddot{y}(t) = u(t)^3$	(b) [ ] $\dot{y}(t) = -3u(t) + t^2$	(c) [x] $\dot{y}(t) = \sqrt{u(t)} + 1$	(d) [ ] $\ddot{y}(t)^2 = u(t)^t + e^{u(t)}$

2. With which of the following formulas you can **NOT** compute the conditional joint distribution of N independent measurements $y_N = (y_1, \dots, y_N)$ given θ? $p(y_N|\theta) \neq \dots$

(a) [ ] $\exp\left(\sum_{i=0}^{N} \ln(p(y(i)|\theta))\right)$	(b) [x] $\int p(y_N|\theta)p(\theta)\mathrm{d}\theta$
(c) [ ] $\prod_i p(y(i)|\theta)$	(d) [ ] $\int p(y_N|x_N,\theta)p(x_N)\mathrm{d}x_N$

3. 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) [x] MAP assumes a linear model.	(d) [ ] $\hat{\theta}_{\mathrm{MAP}} = \arg\min_{\theta \in \mathbb{R}} [-\log(p(y_N|\theta)) - \log(p(\theta))]$

4. You are given a pendulum which is by nature a nonlinear 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 should you use to estimate the parameters θ?

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

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 continuous and infinite flow of measurement data?

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

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

7. We would like to know the unknown probability θ that a phone does **NOT** break 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 23 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the maximum likelihood (ML) estimate of θ ?

(a) [ ] $23\log\theta + 77\log(1-\theta)$	(b) [ ] $\log(23\theta) + \log(77(1-\theta))$
(c) [ ] $-\log(77\theta) - \log(23(1-\theta))$	(d) [x] $-77\log\theta - 23\log(1-\theta)$

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 and c are correct!]

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

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

Points on page (max. 8) | |

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

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) [x] $y(k) = y(k-1)\theta_1 + \sqrt{\theta_2 u(k)}$ | (b) [ ] $y(k) = \theta_1 \sqrt{u(k)}$

(c) [ ] $y(k) = \exp(y(k-1)) \cdot (\theta_1 + \theta_2 u(k))$ | (d) [ ] $y(k) = \theta_1 u(k)^2 + \theta_2 \exp(u(k))$

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

(a) [x] $u(k) - \sqrt{\pi} u(k-2)$ | (b) [ ] $u(k+1) + y(k)$ | (c) [ ] $u(k) \cdot y(k)$ | (d) [ ] $u(k) + \sin(k \cdot \pi)^2$

12. Let θ_{R} denote the *regularized* LLS estimator using L_2 regularization. Which of the following is **NOT** true?

(a) [ ] θ_{R} can be computed analytically. | (b) [ ] θ_{R} is biased.

(c) [ ] θ_{R} incorporates prior knowledge about θ. | (d) [x] θ_{R} is asymptotically biased.

13. We use the Gauss-Newton (GN) algorithm to solve a nonlinear estimation problem. Which of the following statements is **NOT** true *in general*?

(a) [ ] The idea of GN is to linearize the residual function. | (b) [ ] The inverse of the GN Hessian approximates Σ_θ.

(c) [ ] GN uses a Hessian approximation. | (d) [x] GN finds the global minimizer of the objective function.

14. Which one of the following statements is **NOT** true for FIR models:

(a) [ ] They are a special class of ARX models | (b) [ ] Output error minimization is a convex problem.

(c) [ ] The output does not depend on previous outputs. | (d) [x] The impulse response is constant.

15. For a discrete time LTI system $x_{k+1} = Ax_k + Bu_k$, $k = 0, 1$, explicitly state the forward simulation map $f_{\mathrm{sim}} : \mathbb{R}^{n_x + 2n_u} \to \mathbb{R}^{3n_x}$, $(x_0, u_0, u_1) \mapsto (x_0, x_1, x_2)$.

$$f(x_0, u_0, u_1) = \begin{bmatrix} x_0 \\ Ax_0 + Bu_0 \\ A^2 x_0 + ABu_0 + Bu_1 \end{bmatrix}$$

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

(a) [ ] L_2, Gaussian | (b) [x] L_1, Laplace | (c) [ ] L_2, Laplace | (d) [ ] L_1, Gaussian

17. Please identify the most general system equation that still is a Auto Regressive Model with Exogenous Inputs (ARX).

(a) [ ] $y(k) = b_0 u(k) + \ldots + b_{n_b} u(k - n_b)$ | (b) [ ] $y(k) = -a_1 y(k-1) - \ldots - a_{n_a} y(k - n_a)$

(c) [x] $a_0 y(k) + a_1 y(k-1) + \ldots + a_{n_a} y(k - n_a) = b_0 u(k) + \ldots + b_{n_b} u(k - n_b)$ | (d) [ ] $a_0 y(k) + a_1 y^2(k-1) + \ldots + a_{n_a} y^{n_a+1}(k - n_a) = b_0 u(k) + \ldots + b_0 u(k - n_b)$

18. Given the probability density function $p_X(x) = \theta x \exp(-\theta x)$, with parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$, which minimisation problem you need to solve to get a ML-estimate of θ? The problem is: $\min\limits_{\theta} \ldots$?

(a) [ ] $-\log\left(\sum_k \|\theta y(k) \exp(-\theta y(k))\|_2^2\right)$ | (b) [ ] $N \log(\theta) + \theta \sum_k y(k)$

(c) [ ] $\sum_k \|\theta y(k) \exp(-\theta y(k))\|_2^2$ | (d) [x] $-N \log(\theta) + \theta \sum_k y(k)$

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) [ ] $\int_{y_N} N\theta_0^{N-2} \exp[-\theta_0 \sum_k y_k] \mathrm{d}y_N$ | (b) [x] θ_0^2 / N

(c) [ ] N/θ_0^2 | (d) [ ] $\int_{y_N} N\theta_0^{N-2} \left(\sum_k y(k)\right) \exp[-\theta_0 \sum_k y_k] \mathrm{d}y_N$

20. Assume you solved a nonlinear parameter estimation problem using the Gauss-Newton algorithm. How can you check that your assumptions on the noise distribution are correct?
Plot the entries of the residual vector, $R(\theta^*)_i$, $i = 1, \ldots, N$, as a histogram and check if it looks like a zero-mean Gaussian.

Points on page (max. 12) | |

Modeling and System Identification – Microexam 2

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg December 19, 2018, 14:10-14:55, 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. For a discrete time LTI system $x_{k+1} = Ax_k + Bu_k$, $k = 0, 1$, explicitly state the forward simulation map $f_{\text{sim}} : \mathbb{R}^{n_x+2n_u} \to \mathbb{R}^{3n_x}$, $(x_0, u_0, u_1) \mapsto (x_0, x_1, x_2)$.

$$f(x_0, u_0, u_1) = \begin{bmatrix} x_0 \\ Ax_0 + Bu_0 \\ A^2x_0 + ABu_0 + Bu_1 \end{bmatrix}$$

2. Please identify the most general system equation that still is a Auto Regressive Model with Exogenous Inputs (ARX).

(a) ☐ $y(k) = b_0u(k) + \ldots + b_{n_b}u(k - n_b)$	(b) ☐ $a_0y(k) + a_1y^2(k-1) + \ldots + a_{n_a}y^{n_a+1}(k-n_a) = b_0u(k) + \ldots + b_0u(k-n_b)$
(c) ☐ $y(k) = -a_1y(k-1) - \ldots - a_{n_a}y(k - n_a)$	(d) ☒ $a_0y(k) + a_1y(k-1) + \ldots + a_{n_a}y(k-n_a) = b_0u(k) + \ldots + b_{n_b}u(k-n_b)$

3. Let θ_{R} denote the *regularized* LLS estimator using L_2 regularization. Which of the following is **NOT** true?

(a) ☒ θ_{R} is asymptotically biased.	(b) ☐ θ_{R} incorporates prior knowledge about θ.
(c) ☐ θ_{R} is biased.	(d) ☐ θ_{R} can be computed analytically.

4. We use the Gauss-Newton (GN) algorithm to solve a nonlinear estimation problem. Which of the following statements is **NOT** true *in general*?

(a) ☒ GN finds the global minimizer of the objective function.	(b) ☐ GN uses a Hessian approximation.
(c) ☐ The inverse of the GN Hessian approximates Σ_θ.	(d) ☐ The idea of GN is to linearize the residual function.

5. Which of the following models is time invariant?

(a) ☐ $\dot{y}(t) = -3u(t) + t^2$	(b) ☐ $t \cdot \ddot{y}(t) = u(t)^3$	(c) ☒ $\dot{y}(t) = \sqrt{u(t)} + 1$	(d) ☐ $\ddot{y}(t)^2 = u(t)^t + e^{u(t)}$

6. With which of the following formulas you can **NOT** compute the conditional joint distribution of N independent measurements $y_N = (y_1, \ldots, y_N)$ given θ? $p(y_N|\theta) \neq \ldots$

(a) ☒ $\int p(y_N|\theta)p(\theta)\mathrm{d}\theta$	(b) ☐ $\prod_i p(y(i)|\theta)$
(c) ☐ $\exp\left(\sum_{i=0}^N \ln(p(y(i)|\theta))\right)$	(d) ☐ $\int p(y_N|x_N, \theta)p(x_N)\mathrm{d}x_N$

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

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

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

(a) ☐ $u(k) + \sin(k \cdot \pi)^2$	(b) ☐ $u(k+1) + y(k)$	(c) ☒ $u(k) - \sqrt{\pi}u(k-2)$	(d) ☐ $u(k) \cdot y(k)$

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

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

10. We would like to know the unknown probability θ that a phone does **NOT** break 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 23 broken smartphones. What is the negative log likelihood function $f(\theta)$ that we need to minimize in order to obtain the maximum likelihood (ML) estimate of θ ?

(a) ☒ $-77\log\theta - 23\log(1-\theta)$	(b) ☐ $23\log\theta + 77\log(1-\theta)$
(c) ☐ $-\log(77\theta) - \log(23(1-\theta))$	(d) ☐ $\log(23\theta) + \log(77(1-\theta))$

Points on page (max. 10)

11. Given the probability density function $p_X(x) = \theta x \exp(-\theta x)$, with parameter θ, and a set of independent measurements $y_N = [y(1), y(2), \ldots, y(N)]^T$, which minimisation problem you need to solve to get a ML-estimate of θ? The problem is: $\min_\theta \ldots$?

(a) ☐ $-\log\left(\sum_k \|\theta y(k)\exp(-\theta y(k))\|_2^2\right)$	(b) ☐ $\sum_k \|\theta y(k)\exp(-\theta y(k))\|_2^2$
(c) ☐ $N\log(\theta) + \theta\sum_k y(k)$	(d) ☒ $-N\log(\theta) + \theta\sum_k y(k)$

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

13. 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 and c are correct!]

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

14. Which one of the following statements is **NOT** true for FIR models:

(a) ☐ They are a special class of ARX models	(b) ☒ The impulse response is constant.
(c) ☐ The output does not depend on previous outputs.	(d) ☐ Output error minimization is a convex problem.

15. You are given a pendulum which is by nature a nonlinear 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 should you use to estimate the parameters θ?

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

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 continuous and infinite flow of measurement data?

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

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 $\Sigma_{\hat{\theta}} \approx \ldots$

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

18. 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

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

20. Assume you solved a nonlinear parameter estimation problem using the Gauss-Newton algorithm. How can you check that your assumptions on the noise distribution are correct?
Plot the entries of the residual vector, $R(\theta^*)_i$, $i = 1, \ldots, N$, as a histogram and check if it looks like a zero-mean Gaussian.

Points on page (max. 10)	