# Modelling and System Identification – Final Exam

Prof. Dr. Moritz Diehl, IMTEK, Universität Freiburg, and OPTEC/ESAT-STADIUS, KU Leuven

March 17, 2015, 14:00-15:30, Freiburg, Georges-Koehler-Allee 101, HS 026&036

| page | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| points on page (max) | 4 | 9 | 8 | 7 | 7 | 8 | 4 | 4 | 0 | 0 |
| points obtained | | | | | | | | | | |
| intermediate sum | | | | | | | | | | |

Note: Klausur eingesehen am: Unterschrift des Prüfers:

Nachname: Vorname: Matrikelnummer:

Fach: Studiengang: Bachelor ☐ Master ☐ Lehramt ☐ Sonstiges ☐

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 sheets (with 4 pages) of notes. Some legal comments are found in a footnote (in German).[1]

1. Give the probability density function (PDF) $p(x)$ for a **normally distributed** random variable with mean $\mu = \mathbf{5}$ and variance $\sigma^2 = \mathbf{9}$ and add a **sketch** including numbers on the $x$-axis:

   $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}$ . **Sketch:**

   | 2 | |
   |---|---|

2. The PDF of a random variable with **Laplace distribution** with mean $\mu$ and variance $\sigma^2 = 2b^2$ (with $b > 0$) is given by the formula

   $$p(x) = C \cdot \exp\left(-\frac{|x-\mu|}{b}\right)$$

   where $C$ is a normalization constant. Compute the value of $C$ on the left and draw a sketch of the Laplace distribution on the right (Tip: for computing $C$ you can use the integral formula $\int_0^\infty e^{-x/b}\mathrm{d}x = b$).

   $C =$ **Sketch**:

   | 2 | |
   |---|---|

| points on page: 4 | |
|---|---|

[1]PRÜFUNGSUNFÄHIGKEIT: Durch den Antritt dieser Prüfung erklaren Sie sich für prüfungsfähig. Sollten Sie sich während der Prüfung nicht prüfungsfahig fühlen, können Sie aus gesundheitlichen Gründen auch während der Prufung von dieser zurücktreten. Gemäß den Prüfungsordnungen sind Sie verpflichtet, die für den Rücktritt oder das Versäumnis geltend gemachten Gründe unverzüglich (innerhalb von 3 Tagen) dem Prüfungsamt durch ein Attest mit der Angabe der Symptome schriftlich anzuzeigen und glaubhaft zu machen. Weitere Informationen: https://www.tf.uni-freiburg.de/studium/pruefungen/pruefungsunfaehigkeit.html.
TÄUSCHUNG/STÖRUNG: Sofern Sie versuchen, während der Prüfung das Ergebnis ihrer Prüfungsleistung durch Täuschung (Abschreiben von Kommilitonen ...) oder Benutzung nicht zugelassener Hilfsmittel (Skript, Buch, Mobiltelefon, ...) zu beeinflussen, wird die betreffende Prüfungsleistung mit "nicht ausreichend" (5,0) und dem Vermerk Täuschung bewertet. Als Versuch gilt bei schriftlichen Prüfungen und Studienleistungen bereits der Besitz nicht zugelassener Hilfsmittel während und nach der Ausgabe der Prüfungsaufgaben. Sollten Sie den ordnungsgemßen Ablauf der Prüfung stören, werden Sie vom Prüfer/Aufsichtsführenden von der Fortsetzung der Prüfung ausgeschlossen. Die Prüfung wird mit "nicht ausreichend" (5,0) mit dem Vermerk Störung bewertet.

3. What does “i.i.d.” stand for?

| 1 | |
|---|---|

4. Regard a random variable $x \in \mathbb{R}^n$ with PDF $p(x)$, and $g$ a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. How is the expectation $\mathbb{E}\{g(x)\}$ defined?

$$\mathbb{E}\{g(x)\} = \dots$$

| 1 | |
|---|---|

5. Regard $m$ random variables $y_1, \dots, y_m$ that are i.i.d., normally distributed with zero mean and unit variance, that form together the random vector $y \in \mathbb{R}^m$. What is the covariance matrix $\Sigma$ of the $n$-dimensional random vector $x$ defined by $\boxed{x = Ay + b}$ (where $A \in \mathbb{R}^{n\times m}$ is a fixed matrix and $b \in \mathbb{R}^n$ a fixed vector )?

| (a) ☐ $\Sigma = bb^\top$ | (b) ☐ $\Sigma = (A^\top A)^{-1}$ | (c) ☐ $\Sigma = AA^\top$ | (d) ☐ $\Sigma = bb^\top + AA^\top$ |
|---|---|---|---|

| 1 | |
|---|---|

6. For the same data as in Question 5, give a formula for the PDF $p(x)$ of the random variable $x \in \mathbb{R}^n$ (Tip: your formula will only contain $x$, $b$ and $\Sigma$, no need to use $A$).

$$p(x) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}} \dots \qquad .$$

| 2 | |
|---|---|

7. Given a sequence of i.i.d. scalar random variables $y(1), \dots, y(N)$, each with mean $\mu$ and variance $\sigma^2$, what is the **variance** $\sigma_m^2$ of the the arithmetic mean $m$ defined by $\boxed{m = \frac{1}{N}\sum_{k=1}^{N} y(k)}$?

$$\sigma_m^2 = \dots$$

| 1 | |
|---|---|

8. Regard three i.i.d. scalar random variables $y(1), y(2), y(3)$, each with mean $\mu$ and variance $\sigma^2$. What is the expectation $\mu_w$ and what the variance $\sigma_w^2$ of the weighted mean $w$ defined by $\boxed{w = \frac{1}{6} \cdot [\, y(1) + 2 \cdot y(2) + 3 \cdot y(3) \,]}$ ?

$$\mu_w = \dots$$

| 1 | |
|---|---|

$$\sigma_w^2 = \dots$$

| 2 | |
|---|---|

| points on page: 9 | |
|---|---|

9. Compare the unweighted mean $m$ and the weighted mean $w$ of the previous two questions for three i.i.d. Gaussian scalar random variables $y(1), y(2), y(3)$ with unknown mean $\mu$ and known variance $\sigma^2$:

(a) Which of the two estimators $m$ and $w$ is a *biased* and which an *unbiased* estimate for the true mean $\mu$?

| | | | |
|---|---|---|---|
| (a) ☐ both are unbiased | (b) ☐ only $w$ is biased | (c) ☐ only $m$ is biased | (d) ☐ both are biased |

| 1 | |
|---|---|

(b) Which of the two estimators has the lower covariance, and why?

| 2 | |
|---|---|

(c) Compute a lower bound on the covariance that any unbiased estimator of $\mu$ can achieve in this situation, and cite the name of the corresponding theorem. Tip: the Fisher information matrix $M$ is defined by $M = \mathbb{E}\{\nabla_\theta^2[-\log p(y|\theta)]\}$.

| 2 | |
|---|---|

10. What is the minimizer $x^*$ of the convex function $f : \mathbb{R} \to \mathbb{R}$, $\boxed{f(x) = 2x^2 + (x-1)^2}$ ?

$\frac{\partial f}{\partial x}(x) =$ $\qquad\qquad$ $x^* =$

| 1 | |
|---|---|

11. Given three scalars $y_1, y_2, y_3 \in \mathbb{R}$, what is the minimizer $\theta^*$ of the weighted least squares function $\boxed{f(\theta) = w_1(y_1 - \theta)^2 + w_2(y_2 - \theta)^2 + w_3(y_3 - \theta)^2}$ with positive weights $w_1, w_2, w_3 > 0$?

$\theta^* = \ldots$

| 2 | |
|---|---|

| points on page: 8 | |
|---|---|

12. Regard the measurement model $y = \Phi\theta + \epsilon$ where $y$ and $\epsilon$ are in $\mathbb{R}^N$ and the components $\epsilon_1, \ldots, \epsilon_N$ of $\epsilon$ are i.i.d. Gaussian with zero mean and unit variance, and $\Phi \in \mathbb{R}^{N \times d}$. Formulate the conditional PDF $p(y|\theta)$ of $y$ given $\theta$ (up to a normalization constant), define the maximum-likelihood estimator $\hat{\theta}_{\mathrm{ML}}$, and give an explicit formula for $\hat{\theta}_{\mathrm{ML}}$ using $\Phi$ and $y$.

$$p(y|\theta) = \mathrm{const} \cdot \ldots$$

| 1 | |
|---|---|

General definition of maximum likelihood estimator

$$\hat{\theta}_{\mathrm{ML}} = \ldots$$

| 1 | |
|---|---|

Explicit formula of ML estimator for the above model:

$$\hat{\theta}_{\mathrm{ML}} = \ldots$$

| 1 | |
|---|---|

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

$$\Psi_N = \ldots$$

| 2 | |
|---|---|

14. Regard a static model $\boxed{y(k) = \theta + \sin(\theta) x(k) + \epsilon(k)}$ with unknown parameter $\theta \in \mathbb{R}$ and i.i.d. Gaussian noise $\epsilon(k)$ with zero mean and unit variance. For a given sequence of $N$ scalar input and output measurements $x(1), \ldots, x(N)$ and $y(1), \ldots, y(N)$, formulate the objective function that the maximum likelihood (ML) estimator $\hat{\theta}_{\mathrm{ML}}$ needs to minimize.

$$\hat{\theta}_{\mathrm{ML}} = \arg\min_{\theta} \ldots$$

| 2 | |
|---|---|

| points on page: 7 | |
|---|---|

15. In the previous question, classify the type of the objective function that the ML estimator needs to minimize.

| | | | |
|---|---|---|---|
| (a) ☐ convex | (b) ☐ linear least squares | (c) ☐ nonlinear least squares | (d) ☐ linear quadratic |

| 1 | |
|---|---|

16. Regard a noisy input-output model $y(k) = p_1 + p_2 x(k)$ with unknown parameters $p_1, p_2$, where the measurements are given by $y_{\mathrm{m}}(k) = y(k) + \epsilon_y(k)$ and $x_{\mathrm{m}}(k) = x(k) + \epsilon_x(k)$, with i.i.d. Gaussian noise with zero mean and unit variance on both inputs $x(k)$ and outputs $y(k)$. Given a sequence of $N$ scalar input and output measurements $x_{\mathrm{m}}(1), \ldots, x_{\mathrm{m}}(N)$ and $y_{\mathrm{m}}(1), \ldots, y_{\mathrm{m}}(N)$, we want to compute the maximum likelihood (ML) estimate for the extended parameter vector $\theta = [p_1, p_2, x(1), \ldots, x(N)]^\top$. Formulate the objective function which the ML estimator $\hat{\theta}_{\mathrm{ML}}$ needs to minimize.

$$\hat{\theta}_{\mathrm{ML}} = \arg\min_{\theta} \ldots$$

| 2 | |
|---|---|

17. Regard an infinite series of measurement data $y(1), y(2), \ldots, y(N), \ldots$ generated by the model equation $y(k) = \theta_0 + \epsilon(k)$ where $\theta_0$ is the true but unknown mean, and $\epsilon(k)$ is i.i.d. zero mean Gaussian noise. We define a simple estimator $\hat{\theta}(N)$ for $\theta$ by just taking the latest sample, i.e. $\hat{\theta}(N) := y(N)$. Check if this estimator is unbiased and/or consistent.

| | |
|---|---|
| (a) ☐ not unbiased and not consistent | (b) ☐ unbiased, but not consistent |
| (c) ☐ not unbiased, but consistent | (d) ☐ unbiased and consistent |

| 1 | |
|---|---|

18. Given an autoregressive (AR) nonlinear dynamic system model $y(k) = \theta_1 + \theta_2 y(k-1) y(k-2) + \epsilon(k)$ with i.i.d. noise $\epsilon(k)$ with zero mean and unknown parameter vector $\theta = (\theta_1, \theta_2)^\top$, and given a sequence of $N$ output measurements $y(1), \ldots, y(N)$, we want to compute the estimate $\hat{\theta}_N$ by minimizing the prediction error function $f(\theta) = \sum_{k=3}^{N} \epsilon(k)^2 = \|y - \Psi\theta\|_2^2$. How do we need to choose the vector $y$ and matrix $\Psi$ ?

$$y = \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix}, \qquad \Psi = \begin{bmatrix} \quad \\ \quad \\ \quad \end{bmatrix}$$

| 2 | |
|---|---|

19. Which system is described by the transfer function $G(s) = \frac{s^2+3s}{s^3+1}$ ?

| | | | |
|---|---|---|---|
| (a) ☐ $\dddot{y} + y = \ddot{u} + 3\dot{u}$ | (b) ☐ $\ddot{y} + 3\dot{y} = \dddot{u} + u$ | (c) ☐ $\ddot{y} + 3y = \dddot{u} + u$ | (d) ☐ $\dddot{y} + y = \dot{u} + 3u$ |

| 1 | |
|---|---|

| points on page: 7 | |
|---|---|

20. Compute the transfer function $G(s) = \frac{Y(s)}{U(s)}$ that describes the general LTI-system in state space form $\dot{x}(t) = Ax(t) + Bu(t),\ y(t) = Cx(t) + Du(t)$.

$G(s) =$

| 2 | |
|---|---|

21. What quantity of a continuous time transfer function $G(s)$ shows the Bode amplitude diagram in doubly logarithmic scale?

| | | | |
|---|---|---|---|
| (a) ☐ $\arg G(j\omega)$ | (b) ☐ $G(j\omega)$ | (c) ☐ $\lvert G(e^{j\omega})\rvert$ | (d) ☐ $\lvert G(j\omega)\rvert$ |

| 1 | |
|---|---|

22. Regard the LTI SISO system described by the transfer function $G(s) = \frac{1}{1+Ts}$ with $T = 1$s. When we inject an infinite sinusoidal input of the form $u(t) = 5\sin(t/T)$ into the system, the output $y(t)$ is also sinusoidal. Compute $y(t)$ explicitly. Tip: think of the relation between the Bode-Diagram and $G(s)$.

$y(t) =$

| 2 | |
|---|---|

23. Regard a sampled periodic signal with sampling time $\Delta t$ that contains $N$ samples $u(k)$, with $k = 0, \ldots, N-1$. We assume $N$ to be even. You apply the discrete fourier transform (DFT) defined, for $m = 0, \ldots, N-1$, by

$$U(m) = \sum_{k=0}^{N-1} e^{-\frac{2\pi j}{N}mk} \cdot u(k)$$

(a) Why can we usually discard half of the samples $U(m)$ obtained by performing a DFT? Because:

| | |
|---|---|
| (a) ☐ we mainly use multisines for identification | (b) ☐ we mainly work with real-valued signals |
| (c) ☐ we do not care about higher frequencies | (d) ☐ they are all zeros |

| 1 | |
|---|---|

(b) What is the Nyquist frequency $f_{\text{Nyquist}}$ [Hz] corresponding to this signal?

$f_{\text{Nyquist}} =$

| 1 | |
|---|---|

(c) For a given signal $u(k)$, you observe that nearly all entries of $U(m)$ are zero, apart from $U(5)$ and $U(N-5)$. What frequency (in [Hz]) does the signal contain?

| 1 | |
|---|---|

| points on page: 8 | |
|---|---|

24. Consider a recursive least squares algorithm to estimate the value of the estimator $\theta$. The model equation is

$$x(t) = \theta \cdot \exp(-at),$$

with the corresponding $N$ measurements $y(k) = x(k\Delta t) + \epsilon(k), k = 0, \ldots, N-1$ where $\epsilon(k) \sim \mathcal{N}(0, 4)$ and is i.i.d. You can assume the following prior knowledge on the estimator: $\theta \sim \mathcal{N}(0, 9)$. Complete the following MATLAB code that implements this recursive least squares algorithm. Please only fill in the three lines preceded by a comment.

```
function [theta_hat] = recursiveLS(y,t_start,t_end)

N = length(y); a = 1;
t_vector = linspace(t_start,t_end,N).';
theta_hat = 0;
% Complete code in the following line:
Q =

for i=1:N
    t = t_vector(i);
    % Complete code in the following line:
    phi =

    Q = Q + phi * phi.';
    % Complete code in the following line:
    theta_hat = theta_hat

end
```

| 2 | |
|---|---|

25. Describe in your own words the possible advantages and disadvantages of a pseudorandom binary noise (PRBN) signal in discrete time for identification of SISO LTI systems. The PRBN input signal $u_{\text{PRBN}}(k)$ is defined as follows:

$$u_{\text{PRBN}}(k) = \begin{cases} 1 & \text{with probability 0.5} \\ -1 & \text{with probability 0.5} \end{cases}$$

Advantages:

Disdvantages:

| 2 | |
|---|---|

| points on page: 4 | |
|---|---|

26. Given a sequence of discrete time input output measurements $u(1), \ldots, u(N)$ and $y(1), \ldots, y(N)$, you want to identify a finite impulse response model (FIR) of order 5 with one sample delay, assuming output errors (OE) that are i.i.d. normally distributed, using the maximum likelihood (ML) approach. First formulate the FIR model and the corresponding optimization problem on paper. Second, write a MATLAB code that defines the data of the optimization problem, then solves it, and finally outputs the parameter vector $\theta \in \mathbb{R}^5$ and the resulting predictions $y_{\text{pred}}(5)$. Complete the missing lines below.

(a) FIR model of order 5 with one sample delay and OE (note: the predictor $y(k)$ is only defined for $k \geq 6$):

$y(k) =$

| 1 | |
|---|---|

(b) ML optimization problem:

$$\hat{\theta} = \arg\min_{\cdots}$$

| 1 | |
|---|---|

(c) MATLAB Code:

```
function [theta, y_pred]=fir_estimator(u,y)
N=length(u);
d=5;
theta=zeros(d,1);
y_pred=zeros(N,1);

%formulate the data for the optimization problem based on 'u' and 'y'










%solve the optimization problem to obtain the parameter 'theta'





%compute the predicted outputs 'y_pred'(first five samples remain just zero)





% plot the result in red and green
plot([1:N],y,'r',[1:N], y_pred, 'g');

end
```

| 2 | |
|---|---|

| points on page: 4 | |
|---|---|

**Leeres Blatt für Zwischenrechnungen**

**Leeres Blatt für Zwischenrechnungen**