PURDUE

ECE 64100

Midterm Exam, November 7, Fall 2025

NAME ______________________________________

PUID ____

Exam instructions:

- A fact sheet is included **at the end of this exam** for your use.
- You have 60 minutes to work the exam.
- This is a closed-book and closed-note exam. You may not use or have access to your book, notes, any supplementary reference, a calculator, or any communication device including a cell-phone or computer.
- You may not communicate with any person other than the official proctor during the exam.

To ensure Gradescope can read your exam:

- Write your full name and PUID above and on the top of every page.
- Answer all questions in the area designated for each problem.
- Write only on the front of the exam pages.
- DO NOT run over to the next question.

Name/PUID: Key

Problem 1.(35pt) Causal and Non-Causal MRFs

Let X_n be a zero-mean 1-D Gaussian AR process indexed by n, and let h_n be the MMSE causal prediction filter and σ_C^2 be the causal prediction variance.

In addition, let g_n be the MMSE non-causal prediction filter with non-causal prediction variance given by σ_{NC}^2.

Problem 1a) Write an expression for the power spectrum $S_X(\omega)$ of the random process in terms of the causal model parameters (σ_C^2, h_n).

Problem 1b) Write an expression for the power spectrum $S_X(\omega)$ of the random process in terms of the noncausal model parameters (σ_{NC}^2, g_n).

Problem 1c) Derive an equation that relates (σ_C^2, h_n) to (σ_{NC}^2, g_n) to by equating the equations of parts a) and b) above.

Problem 1d) Determine g_n the non-causal prediction filter in terms of h_n, σ_C^2, and σ_{NC}^2.

Problem 1e) Determine σ_{NC}^2 the non-causal prediction variance in terms of (σ_C^2, h_n).

Solution:

Q1a:

$$S_X(\omega) = \frac{\sigma_C^2}{|1 - H(\omega)|^2} \ ,$$

where

$$H(\omega) = \sum_{n=-\infty}^{\infty} h_n \, e^{-j\omega n}$$

Q1b:

$$S_X(\omega) = \frac{\sigma_{NC}^2}{1 - G(\omega)} \ ,$$

where

$$H(\omega) = \sum_{n=-\infty}^{\infty} g_n \, e^{-j\omega n}$$

Q1c:

$$\frac{\sigma_C^2}{|1-H(\omega)|^2} = \frac{\sigma_{NC}^2}{1-G(\omega)}$$

$$\sigma_C^2(1-G(\omega)) = \sigma_{NC}^2|1-H(\omega)|^2$$
$$\sigma_C^2(\delta_n - g_n) = \sigma_{NC}^2(\delta_n - h_n) * (\delta_n - h_{-n})$$

Q1d:

$$\sigma_C^2(\delta_n - g_n) = \sigma_{NC}^2(\delta_n - h_n) * (\delta_n - h_{-n})$$
$$(\delta_n - g_n) = \frac{\sigma_{NC}^2}{\sigma_C^2}(\delta_n - h_n) * (\delta_n - h_{-n})$$
$$g_n = \delta_n - \frac{\sigma_{NC}^2}{\sigma_C^2}(\delta_n - h_n) * (\delta_n - h_{-n})$$

Q1e:

$$g_n\big|_{n=0} = \delta_n - \frac{\sigma_{NC}^2}{\sigma_C^2}(\delta_n - h_n) * (\delta_n - h_{-n})\bigg|_{n=0}$$
$$0 = 1 - \frac{\sigma_{NC}^2}{\sigma_C^2}\left(1 + \sum_{n=1}^{\infty} h_n^2\right)$$
$$\sigma_{NC}^2 = \frac{\sigma_C^2}{(1 + \sum_{n=1}^{\infty} h_n^2)}$$

Name/PUID: ______________________________________

Problem 2.(21pt) Shrinkage Operator

Consider the proximal map given by

$$S_\lambda(y) = \arg \min_{x \in \Re^N} \left\{ \lambda \|x\|_1 + \frac{1}{2}\|x - y\|^2 \right\}$$

Problem 2a) Calculate an explicit form for the function $S_\lambda(y)$ when $N = 1$.

Problem 2b) Calculate an explicit form for the function $S_\lambda(y)$ when $N > 1$.

Problem 2c) Explain in words (i.e., emotionally) what $S_\lambda(y)$ does.

Solution:

Q2a:

$$S_\lambda(y) = \begin{cases} y - \lambda & \text{for } y \geq \lambda \\ 0 & \text{for } |y| < \lambda \\ y + \lambda & \text{for } y \leq \lambda \end{cases}$$

Q2b:

When $N > 1$, then $S_\lambda(y)$ applies the function to each component of y. So we have that

$$S_\lambda(y) = \begin{bmatrix} S_\lambda(y_0) \\ S_\lambda(y_1) \\ \vdots \\ S_\lambda(y_{N-1}) \end{bmatrix} ,$$

where

$$S_\lambda(y) = \begin{cases} y - \lambda & \text{for } y \geq \lambda \\ 0 & \text{for } |y| < \lambda \\ y + \lambda & \text{for } y \leq \lambda \end{cases}$$

Q2c:

$S_\lambda(y)$ sets any value of y that has magnitude less than λ to zero, but it allows values with magnitude large than λ to be maintained but with a value shift towards 0.

Name/PUID: ________________________________

Problem 3.(21pt) Proximal Maps

Consider the proximal map given by

$$H(y) = \arg\min_{x \in \Re^N} \left\{ \frac{1}{2\sigma^2} \|y - x\|^2 + h(x) \right\}$$

For this problem, we will interpret $H(y)$ as a MAP estimate of $\hat{x}$ given y.

Problem 3a) What is the forward model for this MAP estimate? Express your answer by giving an expression for Y given X.

Problem 3b) What is the prior model for this MAP estimate? Express your answer by giving an expression for $p(x)$.

Problem 3c) What happens if you iterate $H(y)$, i.e., you do the following:

$$\text{Repeat}\{x \leftarrow H(x)\}$$

Problem 3d) Imagine that you would like to learn the proximal MAP $H_\theta(y)$ from training data. Then how would you generate the training data, and how would you estimate θ?

Solution:

Q3a:

$$Y = X + \sigma W \ ,$$

where $W \sim N(0, I)$.

Q3b:

$$p(x) = \frac{1}{Z} \exp\{-h(x)\}$$

Q3c:

You should converge to the most probable value of x given by

$$x^* = \arg\min \{h(x)\}$$

Q3d:

You would first generate K samples from the prior distribution $\{X_k\}_{k=0}^{K-1}$. Then for each sample, you should generate a corresponding image Y_k with independent additive white Gaussian noise (AWGN).

$$Y_k = X_k + \sigma W_k \ ,$$

where $W_k \sim N(0, I)$.

Then to design an approximate proximal map, H_{θ^*} you would select the parameter θ^* so that it minimizes a loss function $L(\theta)$ so that

$$\theta^* = \arg\min_{\theta} L(\theta) \ ,$$

where

$$L(\theta) = \frac{1}{K} \sum_{k=0}^{K-1} \| X_k - H_\theta(Y_k) \|^2 \ .$$

So then H_{θ^*} is a denoiser that minimizes the total squared error on the training data set.

Name/PUID: ______________________________

Problem 4.(35pt) Contraction Mappings

Consider a function $H : \Re^2 \to \Re^2$ given by $y = H(x)$ where

$$H(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x$$

Problem 4a) Is $H(x)$ a contraction map?

Problem 4b) Is $H(x)$ non-expansive?

Problem 4c) Does the following iteration converge?

$$\text{Repeat}\{x \leftarrow H(x)\}$$

Justify your answer.

Problem 4d) Does the following iteration converge?

$$\text{Repeat}\{x \leftarrow (1-\rho)x + \rho H(x)\} \quad \text{for } \rho \in (0,1)$$

Justify your answer.

Problem 4e) What does the iteration of 4d converge to?

Solution:

Q4a:

No

Q4b:

Yes

Q4c:

No because it will just reflect about the diagonal with each iteration.

Q4d:

Yes, because this is a Mann iteration and $H(x)$ has a fixed point for $x = [t, t]$ for any t. So the Mann iteration must converge to one of these fixed points.

Q4e:

If $x^0 \in \Re^2$ is the initial value, and x^∞ is the asymptotic value, then the iterations will converge to the projection of x onto the diagonal. So that results in

$$\begin{aligned} x^\infty &= \frac{< \mathbf{1}, x^0 > \mathbf{1}}{\sqrt{2}} \\ &= \left[\frac{x_0^0 + x_0^0}{2}, \frac{x_0^0 + x_0^0}{2} \right] . \end{aligned}$$

ECE641 Fact Sheet

Maximum Likelihood (ML) Estimator (Frequentist)

$$\hat{\theta} = \arg\max_{\theta\in\Omega} p_\theta(Y) = \arg\max_{\theta\in\Omega} \log p_\theta(Y)$$
$$0 = \nabla_\theta p_\theta(Y)|_{\theta=\hat{\theta}}$$
$$\hat{\theta} = T(Y)$$
$$\bar{\theta} = \mathbb{E}_\theta[\hat{\theta}]$$
$$\text{bias}_\theta = \bar{\theta} - \theta \qquad \text{var}_\theta = \mathbb{E}_\theta[(\hat{\theta} - \bar{\theta})^2]$$
$$\text{MSE} = \mathbb{E}_\theta[(\hat{\theta} - \theta)^2] = \text{var}_\theta + (\text{bias}_\theta)^2$$

For $Y = AX + W$, where X and W are independent zero mean Gaussian distributed with R_X and R_W, respectively. Then the ML estimate is find by maximizing $log(p_{y/x}(y/x))$:

$$\hat{X}_{ML} = (A^t R_W^{-1} A)^{-1} A^t R_W^{-1} y$$

Maximum A Posteriori (MAP) Estimator

$$\begin{aligned}\hat{X}_{MAP} &= \arg\max_{x\in\Omega} p_{x|y}(x|Y) \\ &= \arg\max_{x\in\Omega} \log p_{x|y}(x|Y) \\ &= \arg\min_{x\in\Omega}\{-\log p_{y|x}(y|x) - \log p_x(x)\}\end{aligned}$$

For $Y = AX + W$, where X and W are independent zero mean Gaussian distributed with R_X and R_W, respectively. Then the MAP or equivalently MMSE estimate is:

$$\hat{X}_{MAP} = (A^t R_W^{-1} A + R_X^{-1})^{-1} A^t R_W^{-1} y$$

Power Spectral Density (zero-mean WSS Gaussian process)

1D DTFT:

$$S_X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} R(n) e^{-j\omega n}$$

2D DSFT:

$$S_X(e^{j\omega_1}, e^{j\omega_2}) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} R(m,n) e^{-j\omega_1 m - j\omega_2 n}$$

Causal Gaussian Models

$\sigma_n^2 \triangleq \mathbb{E}[\mathcal{E}_n^2]$, $\hat{X} = HX$, $\mathcal{E} = (I - H)X = AX$, $\mathbb{E}[\mathcal{E}\mathcal{E}^t] = \Lambda$, $\Lambda = \text{diag}\{\sigma_1^2, \sigma_2^2, ..., \sigma_N^2\}$
$p_x(x) = |det(A)| p_{\mathcal{E}}(Ax)$, $|det(A)| = 1$,
$R_X = (A^t \Lambda^{-1} A)^{-1}$

1-D Gaussian AR models:

- Toeplitz $H_{i,j} = h_{i-j}$
- Circulant $H_{i,j} = h_{(i-j) mod N}$
- P^{th} order IIR filter $X_n = \mathcal{E}_n + \sum_{i=1}^{P} X_{n-i} h_i$, $R_{\mathcal{E}}(i-j) = \mathbb{E}[\mathcal{E}_i \mathcal{E}_j] = \sigma_c^2 \delta_{i-j}$
- $R_X(n) * (\delta_n - h_n) * (\delta_n - h_{-n}) = R_{\mathcal{E}}(n) = \sigma_c^2 \delta_n$, $S_X = \frac{\sigma_c^2}{|1-H(\omega)|^2}$

2-D Gaussian AR:

- $\mathcal{E}_s = X_s - \sum_{r\in W_p} h_r X_{s-r}$,
- Toeplitz block Toeplitz $H_{mN+k,nN+l} = h_{m-n,k-l}$

Non-causal Gaussian Models

- $\sigma_n^2 \triangleq \mathbb{E}[\mathcal{E}_n^2 | X_i, i \neq n]$, $B_{i,j} = \frac{1}{\sigma_i^2}(\delta_{i-j} - g_{i,j})$, $\sigma_n^2 = (B_{n,n})^{-1}$, $g_{n,i} = \delta_{n-i} - \sigma_n^2 B_{n,i}$ (homogeneous: $g_{i,j} = g_{i-j}, \sigma_i^2 = \sigma_{NC}^2$)
- $G_{i,j} = g_{i,j}$, $\Gamma = \text{diag}\{\sigma_1^2, \sigma_2^2, ..., \sigma_N^2\}$, $B = \Gamma^{-1}(I - G)$, $\Gamma = \text{diag}(B)^{-1}$, $G = I - \Gamma B$, $\mathbb{E}[\mathcal{E}_n X_{n+k}] = \sigma_{NC}^2 \delta_k$
- $R_X(n) * (\delta_n - g_n) * (\delta_n - g_{-n}) = R_{\mathcal{E}}(n) = \sigma_{NC}^2(\delta_n - g_n)$, $S_X = \frac{\sigma_{NC}^2}{1-G(\omega)}$, $R_X(n) * (\delta_n - g_n) = \sigma_{NC}^2 \delta_n$
- **Relationship b/w AR and GMRF:** $\sigma_{NC}^2 = \frac{\sigma_c^2}{1+\sum_{n=1}^{P} h_n^2}$, $g_n = \delta_n - \frac{(\delta_n - h_n)*(\delta_n - h_{-n})}{1+\sum_{n=1}^{P} h_n^2} (= \frac{\rho}{1+\rho^2}(\delta_{n-1} + \delta_{n+1}),\ P = 1)$

Surrogate Function

Our objective is to find a surrogate function $\rho(\Delta; \Delta')$, to the potential function $\rho(\Delta)$.

Maximum Curvature Method

Assume the surrogate function of the form

$$\rho(\Delta; \Delta') = \alpha_1 \Delta + \frac{\alpha_2}{2}(\Delta - \Delta')^2$$

where $\alpha_1 = \rho'(\Delta')$ and $\alpha_2 = \max_{\Delta\in\mathbb{R}} \rho''(\Delta)$.

Symmetric Bound Method

Assume that potential function is bounded by symmetric and quadratic function of Δ, then the surrogate function is

$$\rho(\Delta;\Delta') = \frac{\alpha_2}{2}\Delta^2$$

which results in the following symmetric bound surrogate function:

$$\rho(\Delta;\Delta') = \begin{cases} \frac{\rho'(\Delta')}{2\Delta'}\Delta^2 & \text{if} \quad \Delta' \neq 0 \\ \frac{\rho''(0)}{2}\Delta^2 & \text{if} \quad \Delta' = 0 \end{cases}$$

Review of Convexity in Optimization

Definition A.6. ***Closed, Bounded, and Compact Sets***

Let $\mathcal{A} \subset \mathbb{R}^N$, then we say that $\mathcal{A}$ is:

- **Closed** if every convergent sequence in $\mathcal{A}$ has its limit in $\mathcal{A}$.
- **Bounded** if $\exists M$ such that $\forall x \in \mathcal{A}$, $\|x\| < M$.
- **Compact** if $\mathcal{A}$ is both closed and bounded.

Definition A.11. ***Closed Functions***

We say that function $f : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ is **closed** if for all $\alpha \in \mathbb{R}$, the sublevel set $\mathcal{A}_\alpha = \{x \in \mathbb{R}^N : f(x) \leq \alpha\}$ is closed set.

Theorem A.6. ***Continuity of Proper, Closed, Convex Functions***

Let $f : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ be a proper convex function. Then f is closed if and only if it is lower semicontinuous.

Optimization Methods:

Gradient Descent: $x^{(k+1)} = x^{(k)} - \beta \nabla f(x^{(k)})$
Gradient Descent with Line Search:
$d^{(k)} = -\nabla f(x^{(k)})$

α solves the equation : $0 = \frac{\partial f(x^{(k)} + \alpha d^{(k)})}{\partial \alpha} = [\nabla f(x^{(k)} + \alpha d^{(k)})]^t d^{(k)}$.

Update: $x^{(k+1)} \leftarrow x^k + \alpha \frac{\|d^{(k)}\|^2}{\|d^{(k)}\|_Q^2} d^{(k)}$ where $Q = A^t \Lambda A + B$
Coordinate Descent :
$\alpha = \frac{(y-Ax)^t \Lambda A_{*,s} - x^t B_{*,s}}{\|A_{*,s}\|_\Lambda^2 + B_{s,s}}$ (for $Y|X \sim N(AX, \Lambda^{-1})$)
$x_s \leftarrow x_s + \frac{(y-Ax)^t A_{*,s} - \lambda(x_s - \Sigma_{r \in \partial s} g_{s-r} x_r)}{\|A_{*,s}\|^2 + \lambda}$, $\lambda = \frac{\sigma^2}{\sigma_x^2}$

Pairwise quadratic form identity

$x^t B x = \sum_{s \in S} a_s x_s^2 + \frac{1}{2} \sum_{s \in S} \sum_{r \in S} b_{s,r} |x_s - x_r|^2$, $a_s = \sum_{r \in S} B_{s,r}$, $b_s = -B_{s,r}$

Miscellaneous

For any invertible matrix A, 1. $\frac{\partial |A|}{\partial A} = |A| A^{-1}$ 2. $\frac{\partial tr(BA)}{\partial A} = B$ 3. $tr(AB) = tr(BA)$