EE 641 Midterm Exam

October 13, Fall 2017

Name: ANSWER Key

Instructions

The following is an in-class closed-book exam.

- This exam contains 3 problems worth a total of 100 points.
- You may not use any notes, textbooks, or calculators.
- You are allowed up to 55 minutes to complete the exam.

Good luck.

Problem 1. (35pt)

Let $X_1, \cdots, X_n$ be i.i.d. Gaussian random vectors with distribution $N(\mu, R)$ where $\mu \in \mathbb{R}^p$ and $R \in \mathbb{R}^{p\times p}$ is a symmetric positive-definite matrix, and let $X = [X_1, \cdots, X_n]$ be the $p \times n$ matrix containing all the random vectors. Let $\theta = [\mu, R]$ denote the parameter vector for the distribution, and let b and S be sufficient statistics defined by

$$b = \sum_{k=1}^{n} X_k \tag{1}$$

$$S = \sum_{k=1}^{n} X_k X_k^t = XX^t \; . \tag{2}$$

a) Derive the following expressions for the probability density of $p(x|\theta)$.

$$p(x|\theta) = \frac{1}{(2\pi)^{np/2}} |R|^{-n/2} \exp\left\{ -\frac{1}{2}\mathrm{tr}\left\{SR^{-1}\right\} + b^t R^{-1}\mu - \frac{n}{2}\mu^t R^{-1}\mu \right\} \tag{3}$$

b) Compute the joint ML estimate of μ and R.

Hints: For any invertible matrix A,

$$\frac{\partial |A|}{\partial A} = |A|A^{-1}$$

$$\frac{\partial \mathrm{tr}(BA)}{\partial A} = B \; ,$$

and for any two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times m}$,

$$\mathrm{tr}\{AB\} = \mathrm{tr}\{BA\} \; .$$

Name: ____________________

a) $$p(x|\theta) = p(x_1|\theta)\,p(x_2|\theta)\cdots p(x_n|\theta)$$

$$= \prod_{i=1}^{n} \frac{1}{(2\pi)^{\frac{p}{2}}} |R|^{-1/2} \exp\left\{-\frac{1}{2}(x_i-\mu)R^{-1}(x_i-\mu)\right\}$$

$$= (2\pi)^{-np/2} |R|^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(x_i^t R^{-1} x_i + \mu^t R^{-1}\mu - 2x_i^t R^{-1}\mu\right)\right\}$$

$$= \frac{1}{(2\pi)^{np/2}} |R|^{-n/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left\{\sum_{i=1}^{n} x_i^t R^{-1} x_i\right\} - \frac{n}{2}\mu^t R^{-1}\mu + \sum_{i=1}^{n} x_i^t R^{-1}\mu\right\}$$

$$= \frac{1}{(2\pi)^{np/2}} |R|^{-n/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\left\{\sum_{i=1}^{n} x_i x_i^t R^{-1}\right\} - \frac{n}{2}\mu^t R^{-1}\mu + \sum_{i=1}^{n} x_i^t R^{-1}\mu\right\}$$

$$= \frac{1}{(2\pi)^{np/2}} |R|^{-n/2} \exp\left\{-\frac{1}{2}\mathrm{tr}\{SR^{-1}\} + b^t R^{-1}\mu - \frac{n}{2}\mu^t R^{-1}\mu\right\} \quad \square$$

b) Let $\hat{\mu}_{ML}, \hat{R}_{ML}$ be ML estimate for the parameters.

$$\hat{\mu}_{ML}, \hat{R}_{ML} = \arg\max_{\mu, R} p(X|\theta)$$

$$= \arg\max_{\mu, R} \log p(X|\theta).$$

$$\log p(x|\theta) = -np/2 \log(2\pi) + \frac{n}{2}\log|R^{-1}| - \frac{1}{2}\mathrm{tr}\{SR^{-1}\} - \frac{n}{2}\mu^t R^{-1}\mu + b^t R^{-1}\mu$$

To compute ML Estimate, we need to solve.

$$\begin{cases} \dfrac{\partial \log p(x|\theta)}{\partial \mu} = R^{-1}b - nR^{-1}\mu = 0. \\ \dfrac{\partial \log p(x|\theta)}{\partial R^{-1}} = \dfrac{n}{2}\dfrac{1}{|R^{-1}|}|R^{-1}|\,R - \dfrac{S}{2} + b^t\mu - \dfrac{n}{2}\mu^t\mu = 0 \end{cases}$$

We have
$$\begin{cases} \hat{\mu} = b/n \\ \hat{R} = S/n - \hat{\mu}\hat{\mu}^t \end{cases}$$

check Hessian Matrix P.D. so.

$$\hat{\mu}_{ML} = b/n$$

$$\hat{R}_{ML} = S/n + \hat{\mu}\hat{\mu}^t \quad \square$$

Problem 2. (35pt)

Let $X \in \mathbb{R}^N$ be an N pixel image that we would like to measure, and let $Y \in \mathbb{R}^N$ be the noisy measurements given by

$$Y = AX + W$$

where A is an $N \times N$ nonsingular matrix, W is a vector of i.i.d. Gaussian noise with $W \sim N(\mathbf{0}, \Lambda^{-1})$. Furthermore, in a Bayesian framework, assume that X is a GMRF with noncausal prediction filter g_s and noncausal prediction variance σ^2.

a) Derive an expression for the ML estimate of X.

b) Derive an expression for the MAP cost function $f(x)$.

c) Derive an expression for the MAP estimate of X under the assumption that X is a zero mean GMRF with inverse covariance matrix B.

a) $P_{Y|X}(Y|x) = \frac{1}{(2\pi)^{N/2}} |\Lambda|^{1/2} \exp\left\{-\frac{1}{2}(Y-AX)^t \Lambda (Y-AX)\right\}$

$\hat{X}_{ML} = \arg\max_x P(Y|X)$ take derivative of $P(Y|X)$ and set it to zero.

$= (A^t \Lambda A)^{-1} A^t \Lambda Y = A^{-1} Y$

b) $\log P_{X|Y}(X|Y) = \log P_{Y|X}(Y|X) + \log P_X(X) - \log P_Y(Y)$.

$= -\frac{1}{2}(Y-AX)^t \Lambda (Y-AX) - \frac{1}{2} X^t B X + c(Y)$ → constant w.r.t X.

Where B is N by N Matrix, and $B_{i,j} = \frac{1}{\sigma^2}(\delta_{i-j} - g_{i-j})$.

Then the MAP cost function $f(x)$ is defined as. (ignoring $c(y)$)

$f(x) = -\log P_{X|Y}(X|Y) + c(Y) = \frac{1}{2}(Y-AX)^t \Lambda (Y-AX) + \frac{1}{2} X^t B X$

c) $X_{MAP} = \arg\min_X f(x)$

$= \arg\min_X \left\{\frac{1}{2}(Y-AX)^t \Lambda (Y-AX) + \frac{1}{2} X^t B X\right\}$ take derivative w.r.t x and set to zero

$= (A^t \Lambda A + B)^{-1} A^t \Lambda Y$

Problem 3. (30pt)
Let $X \in \mathbb{R}^N$ have a Gibbs distribution with the form

$$p(x|\sigma) = \frac{1}{z} \exp\left\{ - \sum_{\{s,r\} \in \mathcal{P}} b_{s,r} \rho\left(\frac{x_s - x_r}{\sigma} \right) \right\} .$$

Then prove that the partition function is given by

$$\begin{aligned} z(\sigma) &= \int_{\mathbb{R}^N} \exp\left\{ - \sum_{\{s,r\} \in \mathcal{P}} b_{s,r} \rho\left(\frac{x_s - x_r}{\sigma} \right) \right\} dx \\ &= z_0 \, \sigma^N , \end{aligned}$$

where

$$z_0 = \int_{\mathbb{R}^N} \exp\left\{ - \sum_{\{s,r\} \in \mathcal{P}} b_{s,r} \rho(x_s - x_r) \right\} dx .$$

Problem 3. (30pt)
Let $X \in \mathbb{R}^N$ have a Gibbs distribution with the form

$$p(x|\sigma) = \frac{1}{z} \exp\left\{ - \sum_{\{s,r\}\in\mathcal{P}} b_{s,r}\rho\left(\frac{x_s - x_r}{\sigma}\right) \right\} .$$

Then prove that the partition function is given by

$$\begin{aligned} z(\sigma) &= \int_{\mathbb{R}^N} \exp\left\{ - \sum_{\{s,r\}\in\mathcal{P}} b_{s,r}\rho\left(\frac{x_s - x_r}{\sigma}\right) \right\} dx \\ &= z_0\, \sigma^N , \end{aligned}$$

where

$$z_0 = \int_{\mathbb{R}^N} \exp\left\{ - \sum_{\{s,r\}\in\mathcal{P}} b_{s,r}\rho(x_s - x_r) \right\} dx .$$

$$Z(\sigma) = \int_{R^N} \exp\{-\sum_{\{s,r\}\in P} b_{s,r}\, \rho(\tfrac{x_s - x_r}{\sigma})\}\, dx$$

$$= \int_{x_1} \cdots \int_{x_N} \exp\{-\sum_{\{s,r\}\in P} b_{s,r}\, \rho(\tfrac{x_s - x_r}{\sigma})\}\, dx_1 \cdots dx_N \quad (\text{Let } \sigma x_s' = x_s)$$

$$\overset{\text{Let}}{=\!=} \int_{x_1'} \cdots \int_{x_N'} \exp\{-\sum_{\{s,r\}\in P} b_{s,r}\, \rho(x_s' - x_r')\, d\sigma x_1 \cdots d\sigma x_N$$

$$= \sigma^N \int_{x_1} \cdots \int_{x_N} \exp\{-\sum_{\{s,r\}\in P} b_{s,r}\, \rho(x_s - x_r)\, dx_1 \cdots dx_N$$

$$= \sigma^N \cdot \int_{R^N} \exp\{-\sum_{\{s,r\}\in P} b_{s,r}\, \rho(x_s - x_r)\}\, dx$$

$$= Z_0 \sigma^N \quad \square$$