EE 641 Midterm Exam

October 23, Fall 2014

Name: **Key** ______________________________

Instructions

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

- This exam contains 3 problems worth a total of 110 points.
- You may not use any notes, textbooks, or calculators.
- You are allowed up to 75 minutes to complete the exam.

Good luck.

Problem 1. (30pt)

Let $X \sim N(0, R)$ where R is a $p \times p$ symmetric positive-definite matrix. Further define the precision matrix, $B = R^{-1}$ and use the notation

$$B = \begin{bmatrix} 1/\sigma^2 & A \\ A^t & C \end{bmatrix} ,$$

where $A \in \mathbb{R}^{1\times(p-1)}$ and $C \in \mathbb{R}^{(p-1)\times(p-1)}$.

a) Calculate the marginal density of X_1, the first component of X.

b) Calculate the conditional density of X_1 given all the remaining components, $Y = [X_2, \cdots, X_p]^t$.

c) What is the conditional mean and covariance of X_1 given Y?

Name: ______________________________

Solution:

a) It is given that X is a Gaussian random vector. Therefore the marginal distribution of X_1 follows a Gaussian distribution with zero mean and variance equal to the first entry of the covariance matrix, denoted using, $\mathbb{E}[X_1 X_1] = R_{1,1}$. Hence,

$$p_{X_1}(x_1) = \frac{1}{\sqrt{2\pi R_{1,1}}} \exp\Big\{ -\frac{1}{2}\Big(\frac{x_1^2}{R_{1,1}}\Big)\Big\}$$

b) In order to find the conditional density of X_1 given Y, we use, $p(x_1|y) = \frac{p(x)}{p(y)}$.
Absorbing the terms that do not depend on x_1 into the partition function, the conditional density can be written as

$$p(x_1|y) \propto \exp\Big\{ -\frac{1}{2} x^t B x\Big\}$$

Expanding using the given inverse covariance matrix and simplifying, we get

$$p(x_1|y) \propto \exp\Big\{ -\frac{1}{2}(x_1^2/\sigma^2 + 2x_1 A y)\Big\}$$

We can now complete the square inside the exponent.

$$p(x_1|y) \propto \exp\Big\{ -\frac{1}{2\sigma^2}(x_1 + \sigma^2 A y)^2\Big\}$$

The previous expression has a Gaussian density form, which enables easy computation of the partition function. So, the partition function is $\frac{1}{\sqrt{2\pi\sigma^2}}$ and the conditional density can be expressed as

$$p(x_1|y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big\{ -\frac{1}{2\sigma^2}(x_1 + \sigma^2 A y)^2\Big\}$$

c) From the conditional density, the conditional mean is given by $\mathbb{E}[X_1|Y] = -\sigma^2 A y$ and the conditional variance is given by $\text{var}[X_1|Y] = \sigma^2$.

Problem 2. (40pt)

Let X_n be a 1-D zero-mean stationary Gaussian AR process with MMSE causal prediction filter given by $h_n = \rho\delta_{n-1}$ and causal prediction variance σ_c^2.

a) Calculate, $S_X(\omega)$, the power spectral density of the random process.

b) Calculate, $R_X(n)$, the time autocorrelation of the random process.

c) Sketch plots $S_X(\omega)$ and $R_X(n)$ for $\rho = 0.95$.

d) Calculate (σ_{NC}^2, g_n) the noncausal prediction variance and the noncausal prediction filter for the equivalent GMRF.

Name: ________________________________

Solution:

a) Let ε_n denote the prediction error,

$$\begin{aligned} \varepsilon_n &= x_n * (\delta_n - h_n) \\ &= x_n - \rho x_{n-1} \end{aligned}$$

the autocovariance of ε_n is

$$R_\varepsilon[n] = \sigma_c^2 \delta_n$$

and the power spectrum of ε is

$$S_\varepsilon(\omega) = \sigma_c^2 .$$

The model can equivalently be represented in the form of,

$$x_n = \varepsilon_n * q_n$$

where the Fourier transform of q_n is given by

$$Q(\omega) = \frac{1}{1 - H(\omega)} = \frac{1}{1 - \rho e^{-j\omega}} ,$$

and therefore

$$q[n] = \rho^n u[n] .$$

So, we have that

$$\begin{aligned} S_x(\omega) &= S_\varepsilon(\omega) Q(\omega) Q^*(\omega) \\ &= \frac{\sigma_c^2}{1 - 2\rho\cos\omega + \rho^2} \\ &= \frac{\sigma_c^2}{1 + \rho^2 - 2\rho\cos\omega} \\ &= \frac{\sigma_c^2}{1+\rho^2} \frac{1}{1 - 2\frac{\rho}{1+\rho^2}\cos\omega} \end{aligned}$$

b) Continued from a),

$$\begin{aligned} R_x[n] &= R_\varepsilon[n] * q[n] * q[-n] \\ &= \sigma_c^2 \delta[n] * q[n] * q[-n] \\ &= \sigma_c^2 q[n] * q[-n] \\ &= \sigma_c^2 \sum_{k=-\infty}^{\infty} \rho^k u[k] \rho^{n+k} u[n+k] \\ &= \frac{\sigma_c^2 \rho^{|n|}}{1 - \rho^2} \end{aligned}$$

c) Sketchs

d) For the equivalent GMRF, we have

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

By evaluating the equation for $n = 0$, we get

$$\sigma_{NC}^2 = \frac{\sigma_c^2}{1+\rho^2}$$

Using this relationship,

$$\begin{aligned} g_n &= \delta_n - \frac{(\delta_n - h_n) * (\delta_n - h_{-n})}{1+\rho^2} \\ &= \delta_n - \left[\frac{-\rho}{1+\rho^2} \delta_{n-1} + \delta_n + \frac{-\rho}{1+\rho^2} \delta_{n+1} \right] \\ &= \frac{\rho}{1+\rho^2} (\delta_{n-1} + \delta_{n+1}) \end{aligned}$$

Problem 3. (40pt)

Consider the optimization problem

$$\hat{x} = \arg \min_{x \in \mathbb{R}^N} \left\{ ||y - Ax||^2_\Lambda + x^t B x \right\}$$

where A is a nonsingular $N \times N$ matrix, B is a positive-definite $N \times N$ matrix, and Λ is a diagonal and positive-definite matrix..

a) Derive a closed form expression for the solution.

b) Calculate an expression for the gradient descent update using step size $\mu \geq 0$.

c) Calculate an expression for the update of gradient descent with line search.

d) Calculate an expression for the coordinate descent update.

Name: ______________________________

Solution:

a)

Define $f(x) = ||y - Ax||_\Lambda^2 + x^t Bx$.

The first derivative of the solution $\hat{x}$ need to be 0.

$$\nabla f(\hat{x}) = -2A^t\Lambda(y - A\hat{x}) + 2B\hat{x} = 0$$

then,

$$\hat{x} = (A^t\Lambda A + B)^{-1}A^t\Lambda y$$

Since the Hessian of $f(x) = 2A^t\Lambda A + 2B$ is positive definite, $\hat{x} = (A^t\Lambda A + B)^{-1}A^t\Lambda y$ is the solution.

b)

For gradient descent update using step size μ,

$$d^{(k)} = -\nabla f(x^{(k)}) = 2A^t\Lambda(y - Ax^{(k)}) - 2Bx^{(k)}$$

the update is

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \mu d_{x^{(k)}} \\ &= x^{(k)} + 2\mu A^t\Lambda(y - Ax^{(k)}) - 2\mu Bx^{(k)} \end{aligned}$$

c)

For gradient descent with line search,

$$d^{(k)} = -\nabla f(x^{(k)}) = 2A^t\Lambda(y - Ax^{(k)}) - 2Bx^{(k)}$$

the step size α is

$$\alpha = \frac{||d^{(k)}||^2}{||d^{(k)}||_Q^2}$$

where $Q = A^t\Lambda A + B$.

The update is

$$\begin{aligned} x^{(k+1)} &= x^{(k)} + \alpha d^{(k)} \\ &= x^{(k)} + \frac{||d^{(k)}||^2}{||d^{(k)}||_Q^2} d^{(k)} \\ &= x^{(k)} + 2\frac{||A^t\Lambda(y - Ax^{(k)}) - Bx^{(k)}||^2}{||A^t\Lambda(y - Ax^{(k)}) - Bx^{(k)}||_{A^t\Lambda A + B}^2} \left(A^t\Lambda(y - Ax^{(k)}) - Bx^{(k)}\right) \end{aligned}$$

d)

The ICD update can be computed by solving the equation

$$0 = \frac{\partial f(x+\alpha\varepsilon_s)}{\partial\alpha} = \left[\nabla f(x+\alpha\varepsilon_s)\right]^t \varepsilon_s \ .$$

For notational simplicity, let $e = y - Ax$, then the gradient term has the form

$$\nabla f(x+\alpha\varepsilon_s) = -2A^t\Lambda(e-\alpha A\varepsilon_s) + 2B(x+\alpha\varepsilon_s) \ .$$

From this we can evaluate the equation

$$\begin{aligned} 0 &= \left[\nabla f(x+\alpha\varepsilon_s)\right]^t \varepsilon_s \\ &= -2(e-\alpha A\varepsilon_s)^t\Lambda A\varepsilon_s + 2(x+\alpha\varepsilon_s)B\varepsilon_s \end{aligned}$$

Rearranging terms results in

$$\begin{aligned} (e-\alpha A\varepsilon_s)^t\Lambda A\varepsilon_s &= (x+\alpha\varepsilon_s)B\varepsilon_s \\ e^t\Lambda A\varepsilon_s - \alpha\varepsilon_s^t A^t\Lambda A\varepsilon_s &= x^t B\varepsilon_s + \alpha\varepsilon_s^t B\varepsilon_s \\ e^t\Lambda A_{*,s} - \alpha||A_{*,s}||_\Lambda &= x^t B_{*,s} + \alpha B_{s,s} \end{aligned}$$

Then solving for α results in

$$\begin{aligned} \alpha &= \frac{e^t\Lambda A_{*,s} + x^t B_{*,s}}{||A_{*,s}||_\Lambda + B_{s,s}} \\ &= \frac{(y-Ax)^t\Lambda A_{*,s} + x^t B_{*,s}}{||A_{*,s}||_\Lambda + B_{s,s}} \end{aligned}$$

So then the ICD update is given by

$$x \leftarrow x + \alpha\varepsilon_s \ .$$