ASEN 5044 Statistical Estimation for Dynamical Systems
Fall 2017
Homework 5

Out: Thursday 10/11/2018 (posted on Canvas)

Due: Thursday 10/18/2018 (Canvas - **no credit for illegible submissions**)

Show all your work and explain your reasoning. ***Do NOT use numerical or symbolic computation tools for any question.***

1. Simon, Problem 2.8. Note: the problem does NOT say that the RVs are Gaussian, so do not make that assumption.

2. Simon, Problem 2.9.

3. Simon, Problem 2.10.

4. Prove the following two results used in lecture to derive the theoretical expectations for the Gaussian sampling experiment where $x \sim \mathcal{N}(\bar{x}, \sigma_x^2)$, $e \sim \mathcal{N}(0, \sigma_e^2)$, and $y = cx + de$:

a. $\text{cov}(X, Y) = E[(x - \bar{x})(y - \bar{y})] = E[XY] - \bar{x}\bar{y}$,

b. $\text{var}(Y) = E[(y - \bar{y})^2] = c^2\sigma_x^2 + d^2\sigma_e^2$.

5. Consider two continuous random variables x and y, where $y = \ln(x)$ and $x > 0$. Derive analytical closed-form expressions for each of the following:

a. $p(y)$ if $p(x) = \mathcal{U}[a, b]$ (i.e. if x has a uniform pdf for $0 < a \leq x \leq b$);

b. $p(y)$ if $p(\frac{1}{x}) = \mathcal{U}[c, d]$ (i.e. if $x^{-1} = \frac{1}{x}$ has a uniform pdf for $0 < c \leq x^{-1} \leq d$);

c. $p(x)$ if $p(y) = \mathcal{U}[l, m]$ (i.e. if y has a uniform pdf for $l \leq y \leq m$);

d. $p(x)$ if $p(y) = \mathcal{N}(\mu_y, \sigma_y^2)$ (i.e. if y has a Gaussian pdf with mean μ_y and variance σ_y^2);

Advanced Questions ***All students are welcome to try any of these for extra credit (only given if all regular problems turned in on time as well). Submit your responses for these questions with rest of your homework, but make sure these are clearly labeled and start on separate pages – indicate in the .pdf file name (per instructions posted on Canvas) and on the front page of your assignment if you answered these questions, so they can be spotted, graded and recorded more easily.***

AQ1. Simon, Problem 2.13. Note: see the Simon textbook errata – the problem statement incorrectly mislabels the random variable V as B.

AQ2. One convenient property of a normally distributed random variable is that a linear transformation of such a variable is also normally distributed. This exercise will briefly explore another approach to deriving this property and then demonstrate it numerically using Monte Carlo simulations.

This problem will require the use of the concept of *moment generating functions* (MGFs). The joint moment generating function $M_{\mathbf{X}}(t)$, if it exists, of a random vector $\mathbf{X} \in \mathbb{R}^n$ is defined as

$$M_{\mathbf{X}}(\mathbf{t}) \triangleq \mathbb{E}[\exp(\mathbf{t}^T\mathbf{X})]$$

It can be shown that two random vectors are identically distributed if their joint MGFs are identical. The variable $\mathbf{t}$ is the domain of the moment generating function and is analogous to the s domain variable for Laplace transforms.

Consider a discrete-time linear time-invariant system of the form

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{G}\mathbf{u}_{k+1}$$

where $\mathbf{x}_k \sim \mathcal{N}(\hat{\mathbf{x}}_k, \hat{\mathbf{P}}_k)$. The mean $\hat{\mathbf{x}}_k = \mathbb{E}\left[\mathbf{x}_k\right]$ may be thought of as an optimal estimate of $\mathbf{x}_k$, and the covariance $\hat{\mathbf{P}}_k = \mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T\right]$ a measure of the uncertainty in the estimate.

a. Assume for now that $\mathbf{u}_{k+1}$ is deterministic. Given that the joint MGF of $\mathbf{x}_k$ is given by

$$M_{\mathbf{x}_k}(\mathbf{t}) = \exp(\mathbf{t}^T\hat{\mathbf{x}}_k + \frac{1}{2}\mathbf{t}^T\hat{\mathbf{P}}_k\mathbf{t})$$

show that the joint MGF of $\mathbf{x}_{k+1}$ is equal to

$$M_{\mathbf{x}_{k+1}}(\mathbf{t}) = \exp\left(\mathbf{t}^T\left(\mathbf{F}\hat{\mathbf{x}}_k + \mathbf{G}\mathbf{u}_{k+1}\right) + \frac{1}{2}\mathbf{t}^T\left(\mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T\right)\mathbf{t}\right)$$

That is, the joint MGF of $\mathbf{x}_{k+1}$ is identical to a normally distributed random vector with mean $\hat{\mathbf{x}}_{k+1} = \mathbf{F}\hat{\mathbf{x}}_k + \mathbf{G}\mathbf{u}_{k+1}$ and covariance $\hat{\mathbf{P}}_{k+1} = \mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T$.

The following properties may prove useful.

$$\begin{aligned} M_{(\mathbf{X}+\mathbf{C})}(\mathbf{t}) &= \exp\left(\mathbf{t}^T\mathbf{C}\right)M_{\mathbf{X}}(\mathbf{t}) \\ M_{\mathbf{A}\mathbf{X}}(\mathbf{t}) &= M_{\mathbf{X}}(\mathbf{A}^T\mathbf{t}) \end{aligned}$$

Here $\mathbf{X} \in \mathbb{R}^n$ is again a random vector, $\mathbf{C} \in \mathbb{R}^n$ is some deterministic vector, and $\mathbf{A} \in^n \mathbb{R}^n$ is a deterministic square matrix.

b. Now assume that $\mathbf{u}_{k+1}$ is a noisy input such that $\mathbf{u}_{k+1} = \bar{\mathbf{u}}_{k+1} + \mathbf{w}_{k+1}$ with $\mathbf{w}_{k+1} \sim \mathcal{N}(\mathbf{0}_{k+1}, \mathbf{Q})$, and $\mathbf{w}_{k+1}$ and $\mathbf{x}_k$ are independent. Show that the joint MGF of $\mathbf{x}_{k+1}$ is equal to

$$M_{\mathbf{x}_{k+1}}(\mathbf{t}) = \exp\left(\mathbf{t}^T\left(\mathbf{F}\hat{\mathbf{x}}_k + \mathbf{G}\mathbf{u}_{k+1}\right) + \frac{1}{2}\mathbf{t}^T\left(\mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T\right)\mathbf{t}\right)$$

That is, the joint MGF of $\mathbf{x}_{k+1}$ is identical to a normally distributed random vector with mean $\hat{\mathbf{x}}_{k+1} = \mathbf{F}\hat{\mathbf{x}}_k + \mathbf{G}\mathbf{u}_{k+1}$ and covariance $\hat{\mathbf{P}}_{k+1} = \mathbf{F}\hat{\mathbf{P}}_k\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T$.

In addition to the properties provided in Part (a), the following property may prove useful: for two independent random vectors $\mathbf{X} \in \mathbb{R}^n$ and $\mathbf{Y} \in \mathbb{R}^n$,

$$M_{(\mathbf{X}+\mathbf{Y})}(\mathbf{t}) = M_{\mathbf{X}}(\mathbf{t})M_{\mathbf{Y}}(\mathbf{t})$$

c. Using definitions of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{x}_k$ from the two-mass three-spring system presented in Lecture 4 with the same masses and spring constants, compute the corresponding $\mathbf{F}$ and $\mathbf{G}$ matrices using a discretization time step of $\Delta t = 1$ second. Next, generate 50000 normally-distributed samples of $\mathbf{x}_k$ with means and variances defined below and compute $\mathbf{x}_{k+1}$ for each sample. Assume $\mathbf{u}_{k+1} = \begin{bmatrix}0 & 0.1\end{bmatrix}^T$ is deterministic. Plot the histograms of each state component and report the sample means and sample variances of each. Use at least 20 bins. Also compute the theoretical means and variances using the results of Part (a). How do the sample means and variances compare to the theoretical means and variances? Note that the theoretical variances for each component of $\hat{\mathbf{x}}_{k+1}$ can be found by extracting the corresponding diagonal elements of $\hat{\mathbf{P}}_{k+1}$.

$$\hat{\mathbf{x}}_k = \begin{bmatrix} 0.1 \\ 0.3 \\ -0.3 \\ -0.8 \end{bmatrix} \qquad \hat{\mathbf{P}}_k = \begin{bmatrix} 0.01 & 0 & 0 & 0 \\ 0 & 0.04 & 0 & 0 \\ 0 & 0 & 0.01 & 0 \\ 0 & 0 & 0 & 0.04 \end{bmatrix}$$

d. For the same two-mass three-spring system and using the same means and covariances for $\mathbf{x}_k$ as in Part (c), but now taking $\mathbf{u}_{k+1} = \begin{bmatrix}0 & 0.1\end{bmatrix}^T$ to be a noisy input with $q_{11} = q_{22} = 0.01$, generate 50000 normally-distributed samples of $\mathbf{x}_k$ and $\mathbf{u}_{k+1}$ and compute $\mathbf{x}_{k+1}$ for each sample. Plot the histograms of each state component and report the sample means and sample variances of each. Also compute the theoretical means and variances using the results of Part (b). How do the sample means and variances compare to the theoretical means and variances?