Section 2
EE278: Introduction to Statistical Signal Processing (Spring 2017)
Gates b03, 4:30-5:20pm, Monday, 04/17

1. Vector-version LLN
The scalar-version LLN can be stated as follows: for i.i.d. RVs $X_1, X_2, \cdots$, let $S_n = \sum_i X_i$. If $E[X_1] < \infty$ and $\text{Var}[X_1] < \infty$, then for any $\epsilon > 0$,
\[
\lim_{n \to \infty} \Pr \left( \frac{1}{n} S_n - E[X_1] > \epsilon \right) = 0.
\]
Now prove the following vector-version LLN: for i.i.d. random vectors $X_1, X_2, \cdots \in \mathbb{R}^k$, let $S_n = \sum_i X_i$. If $E[X_1] < \infty$ and $\text{Var}[X_1] < \infty$ (the inequality is element-wise), then for any $\epsilon > 0$,
\[
\lim_{n \to \infty} \Pr \left( \max_{j=1,\ldots,k} \left| \frac{1}{n} (S_n)_j - E[(X_1)_j] \right| > \epsilon \right) = 0.
\]
(Hint: use the union bound)

Solution: As $X_1, X_2, \cdots$ are i.i.d. random vectors, then for their $j$-th elements, $(X_1)_j, (X_2)_j, \cdots$ are i.i.d. random variables. Hence the scalar-version LLN applies. In other word, for any $j = 1, \ldots, k$ and any $\epsilon > 0$, we have
\[
\lim_{n \to \infty} \Pr \left( \left| \frac{1}{n} (S_n)_j - E[(X_1)_j] \right| > \epsilon \right) = 0.
\]
Then for any $\epsilon > 0$,
\[
\lim_{n \to \infty} \Pr \left( \max_{j=1,\ldots,k} \left| \frac{1}{n} (S_n)_j - E[(X_1)_j] \right| > \epsilon \right) = \lim_{n \to \infty} \Pr \left( \bigcup_j \left| \frac{1}{n} (S_n)_j - E[(X_1)_j] \right| > \epsilon \right) \leq \lim_{n \to \infty} \sum_{j=1}^k \Pr \left( \left| \frac{1}{n} (S_n)_j - E[(X_1)_j] \right| > \epsilon \right) \leq \sum_{j=1}^k \lim_{n \to \infty} \Pr \left( \left| \frac{1}{n} (S_n)_j - E[(X_1)_j] \right| > \epsilon \right) = 0.
\]

2. Unbiased and consistent estimation of the variance
Let $X_1, X_2, \cdots$ be i.i.d. RVs with mean $\mu$ and variance $\theta$. Let $\hat{\theta}_n$ be an estimator of the variance. We say $\hat{\theta}_n$ is unbiased if $E[\hat{\theta}_n] = \theta$ and $\hat{\theta}_n$ is consistent if $\hat{\theta}_n$ converges to $\theta$ in probability.

(a) If we know the $\mu$, propose an unbiased and consistent estimator of the variance.

(b) If we do not know the variance $\mu$, propose an unbiased and consistent estimate of the variance.

Solution: (a) If we know $\mu$, then the estimator can simply be
\[
\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2.
\]
It is unbiased because
\[
E[\hat{\theta}_n] = E\left[ \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 \right] = \frac{1}{n} \sum_{i=1}^n E[(X_i - \mu)^2] = \theta.
\]

To prove consistence, we note that the terms $(X_i - \mu)^2$ are i.i.d. and then we can use LLN.
(b) We start by substituting $\mu$ in 1 by its empirical estimate $\frac{1}{n} \sum_{i=1}^{n} X_i$. Then the estimator becomes

$$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - \frac{1}{n} \sum_{j=1}^{n} X_j)^2.$$ 

We next examining its expectation:

$$\mathbb{E}[\hat{\theta}_n] = \mathbb{E}[\frac{1}{n} \sum_{i=1}^{n} (X_i - \frac{1}{n} \sum_{j=1}^{n} X_j)^2] = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^{n} \left( X_i^2 - 2X_i \frac{1}{n} \sum_{j=1}^{n} X_j + \left( \frac{1}{n} \sum_{j=1}^{n} X_j \right)^2 \right) \right]$$

$$= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i^2] - 2\frac{1}{n} \sum_{i=1}^{n} X_i \frac{1}{n} \sum_{j=1}^{n} X_j + \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^{n} \left( \frac{1}{n} \sum_{j=1}^{n} X_j \right)^2 \right]$$

$$= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i^2] - \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} X_i X_j = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[X_i^2] - \frac{1}{n^2} \sum_{i \neq j} \mathbb{E}[X_i X_j]$$

$$= \frac{n-1}{n^2} \sum_{i=1}^{n} \mathbb{E}[X_i^2] - \frac{1}{n^2} \sum_{i \neq j} \mathbb{E}[X_i X_j]$$

We know that $\mathbb{E}[X_i^2] = \theta + \mu^2$ and $\mathbb{E}[X_i X_j] = \mu^2$. Then

$$\mathbb{E}[\hat{\theta}_n] = \frac{n-1}{n^2} \sum_{i=1}^{n} (\theta + \mu^2) - \frac{1}{n^2} \sum_{i \neq j} \mu^2 = \frac{n-1}{n^2} n(\theta + \mu^2) - \frac{1}{n^2} n(n-1) \sum_{i \neq j} \mu^2 = \frac{n-1}{n} \theta.$$ 

Then we know $\mathbb{E}[\frac{n}{n-1} \hat{\theta}_n] = \theta$. Therefore our unbiased estimator for the variance should be

$$\hat{\theta}_{n}^{ub} = \frac{n}{n-1} \hat{\theta}_n = \frac{1}{n-1} (X_i - \frac{1}{n} \sum_{j=1}^{n} X_j)^2.$$ 

To show that it is consistent, note that by previous derivation, 

$$\hat{\theta}_{n}^{ub} = \frac{n}{n-1} \left( \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \frac{1}{n} \sum_{i=1}^{n} X_i \frac{1}{n} \sum_{j=1}^{n} X_j \right)$$

By LLN,

$$\frac{1}{n} \sum_{i=1}^{n} X_i \overset{p}{\to} \mu, \quad \frac{1}{n} \sum_{i=1}^{n} X_i^2 \overset{p}{\to} \theta + \mu^2.$$ 

Also, 

$$\frac{n}{n-1} \to 1.$$ 

Then,

$$\hat{\theta}_{n}^{ub} = \frac{n}{n-1} \left( \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \frac{1}{n} \sum_{i=1}^{n} X_i \frac{1}{n} \sum_{j=1}^{n} X_j \right) \overset{p}{\to} 1 (\theta + \mu^2 - \mu^2) = \theta.$$ 

□
3. Covariance matrices

Which of the following matrices can be a covariance matrix? Justify your answer. Either construct a random vector \( X \) with the given covariance matrix as a function of the i.i.d. zero mean unit variance random variables \( Z_1, Z_2, Z_3 \), or establish a contradiction as was done in lecture.

(a) \[
\begin{bmatrix}
1 & 2 \\
0 & 2
\end{bmatrix}
\]
(b) \[
\begin{bmatrix}
2 & 1 \\
1 & 2
\end{bmatrix}
\]
(c) \[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 2 & 2 \\
1 & 2 & 3
\end{bmatrix}
\]
(d) \[
\begin{bmatrix}
1 & 1 & 2 \\
1 & 2 & 3 \\
2 & 3 & 3
\end{bmatrix}
\]

Solution:

a) No: not symmetric.
b) Yes: covariance matrix of \( X_1 = Z_1 + Z_2 \) and \( X_2 = Z_1 + Z_3 \).
c) Yes: covariance matrix of \( X_1 = Z_1 \), \( X_2 = Z_1 + Z_2 \), and \( X_3 = Z_1 + Z_2 + Z_3 \).
d) No: several justifications.
   - \( \sigma_{23}^2 = 9 > \sigma_{22}\sigma_{33} = 6 \), which contradicts the Schwarz inequality.
   - The matrix is not nonnegative definite since the determinant is \(-2\).
   - One of the eigenvalues is negative (\( \lambda_1 = -0.8056 \)).

4. Fun with Gaussian vectors.

Let \( X \) be a Gaussian random vector with mean \( \mu \) and covariance matrix \( K \) given by

\[
\mu = \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} \quad \text{and} \quad \Sigma = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix}.
\]

1. Find the pdf of \( X_1 \).
2. Find the pdf of \( X_2 + X_3 \).
3. Find the pdf of \( 2X_1 + X_2 + X_3 \).
4. Find \( P\{2X_1 + X_2 + X_3 < 0\} \).
5. Find the joint pdf of \( Y = AX \), where \( A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \).

Solution:

1. The marginal pdfs of a jointly Gaussian pdf are Gaussian. Therefore \( X_1 \sim \mathcal{N}(1,1) \).
2. Since \( X_2 \) and \( X_3 \) are independent (\( \sigma_{23} = 0 \)), the variance of the sum is the sum of the variances. Also the sum of two jointly Gaussian random variables is also Gaussian. Therefore \( X_2 + X_3 \sim \mathcal{N}(7, 13) \).
3. Since \( 2X_1 + X_2 + X_3 \) is a linear transformation of a Gaussian random vector,

\[
2X_1 + X_2 + X_3 = \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix},
\]

it is a Gaussian random vector with mean and variance

\[
\mu = \begin{bmatrix} 2 & 1 & 1 \\ 5 & 2 \end{bmatrix} = 9 \quad \text{and} \quad \sigma^2 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = 21.
\]

Thus \( 2X_1 + X_2 + X_3 \sim \mathcal{N}(9, 21) \).
4. Let $Y = 2X_1 + X_2 + X_3$. In part (c) we found that $Y \sim \mathcal{N}(9, 21)$. Thus

$$
\mathbb{P}\{Y < 0\} = \Phi\left(\frac{0 - \mu_Y}{\sigma_Y}\right) = \Phi\left(-\frac{9}{\sqrt{21}}\right) = \Phi(-1.96) = Q(1.96) = 2.48 \times 10^{-2}.
$$

5. In general, $AX \sim \mathcal{N}(A\mu_X, A\Sigma_XA^T)$. For this problem,

$$
\mu_Y = A\mu_X = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ -2 \end{bmatrix}
$$

$$
\Sigma_Y = A\Sigma_XA^T = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 21 & 6 \\ 6 & 12 \end{bmatrix}
$$

Thus $Y \sim \mathcal{N}\left(\begin{bmatrix} 9 \\ -2 \end{bmatrix}, \begin{bmatrix} 21 & 6 \\ 6 & 12 \end{bmatrix}\right)$.