Problem 8.1.5

Since J_1, J_2 and J_3 are independent, we can write

$$P_{\mathbf{K}}(\mathbf{k}) = P_{J_1}(k_1)\, P_{J_2}(k_2 - k_1)\, P_{J_3}(k_3 - k_2)\,. \tag{1}$$

Since $P_{J_i}(j) > 0$ only for integers $j > 0$, we have that $P_{\mathbf{K}}(\mathbf{k}) > 0$ only for $0 < k_1 < k_2 < k_3$; otherwise $P_{\mathbf{K}}(\mathbf{k}) = 0$. Finally, for $0 < k_1 < k_2 < k_3$,

$$\begin{aligned} P_{\mathbf{K}}(\mathbf{k}) &= (1-p)^{k_1-1}p(1-p)^{k_2-k_1-1}p(1-p)^{k_3-k_2-1}p \\ &= (1-p)^{k_3-3}p^3. \end{aligned} \tag{2}$$

Problem 8.2.5

We find the marginal PDFs using Theorem 5.26. First we note that for $x < 0$, $f_{X_i}(x) = 0$. For $x_1 \geq 0$,

$$f_{X_1}(x_1) = \int_{x_1}^{\infty} \left(\int_{x_2}^{\infty} e^{-x_3}\, dx_3 \right) dx_2 = \int_{x_1}^{\infty} e^{-x_2}\, dx_2 = e^{-x_1}. \tag{1}$$

Similarly, for $x_2 \geq 0$, X_2 has marginal PDF

$$f_{X_2}(x_2) = \int_{0}^{x_2} \left(\int_{x_2}^{\infty} e^{-x_3}\, dx_3 \right) dx_1 = \int_{0}^{x_2} e^{-x_2}\, dx_1 = x_2 e^{-x_2}. \tag{2}$$

Lastly,

$$\begin{aligned} f_{X_3}(x_3) &= \int_{0}^{x_3} \left(\int_{x_1}^{x_3} e^{-x_3}\, dx_2 \right) dx_1 \\ &= \int_{0}^{x_3} (x_3 - x_1) e^{-x_3}\, dx_1 \\ &= -\frac{1}{2}(x_3 - x_1)^2 e^{-x_3} \Big|_{x_1=0}^{x_1=x_3} = \frac{1}{2} x_3^2 e^{-x_3}. \end{aligned} \tag{3}$$

The complete expressions for the three marginal PDFs are

$$f_{X_1}(x_1) = \begin{cases} e^{-x_1} & x_1 \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

$$f_{X_2}(x_2) = \begin{cases} x_2 e^{-x_2} & x_2 \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

$$f_{X_3}(x_3) = \begin{cases} (1/2) x_3^2 e^{-x_3} & x_3 \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

In fact, each X_i is an Erlang $(n, \lambda) = (i, 1)$ random variable.

Problem 8.3.1

For discrete random vectors, it is true in general that

$$P_{\mathbf{Y}}(\mathbf{y}) = \mathrm{P}\left[\mathbf{Y} = \mathbf{y}\right] = \mathrm{P}\left[\mathbf{AX} + \mathbf{b} = \mathbf{y}\right] = \mathrm{P}\left[\mathbf{AX} = \mathbf{y} - \mathbf{b}\right]. \qquad (1)$$

For an arbitrary matrix $\mathbf{A}$, the system of equations $\mathbf{Ax} = \mathbf{y} - \mathbf{b}$ may have no solutions (if the columns of $\mathbf{A}$ do not span the vector space), multiple solutions (if the columns of $\mathbf{A}$ are linearly dependent), or, when $\mathbf{A}$ is invertible, exactly one solution. In the invertible case,

$$P_{\mathbf{Y}}(\mathbf{y}) = \mathrm{P}\left[\mathbf{AX} = \mathbf{y} - \mathbf{b}\right] = \mathrm{P}\left[\mathbf{X} = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})\right] = P_{\mathbf{X}}\left(\mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})\right). \quad (2)$$

As an aside, we note that when $\mathbf{Ax} = \mathbf{y} - \mathbf{b}$ has multiple solutions, we would need to do some bookkeeping to add up the probabilities $P_{\mathbf{X}}(\mathbf{x})$ for all vectors $\mathbf{x}$ satisfying $\mathbf{Ax} = \mathbf{y} - \mathbf{b}$. This can get disagreeably complicated.

Problem 8.4.2

The mean value of a sum of random variables is always the sum of their individual means.

$$\mathrm{E}\left[Y\right] = \sum_{i=1}^{n} \mathrm{E}\left[X_i\right] = 0. \qquad (1)$$

The variance of any sum of random variables can be expressed in terms of the individual variances and co-variances. Since the $\mathrm{E}[Y]$ is zero, $\mathrm{Var}[Y] = \mathrm{E}[Y^2]$. Thus,

$$\begin{aligned}
\mathrm{Var}[Y] &= \mathrm{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\right] \\
&= \sum_{i=1}^{n} \mathrm{E}\left[X_i^2\right] + \sum_{i=1}^{n}\sum_{j\neq i} \mathrm{E}\left[X_i X_j\right]. \qquad (2)
\end{aligned}$$

Since $\mathrm{E}[X_i] = 0$, $\mathrm{E}[X_i^2] = \mathrm{Var}[X_i] = 1$ and for $i \neq j$,

$$\mathrm{E}\left[X_i X_j\right] = \mathrm{Cov}\left[X_i, X_j\right] = \rho \qquad (3)$$

Thus, $\mathrm{Var}[Y] = n + n(n-1)\rho$.

Problem 8.4.12

Given an arbitrary random vector $\mathbf{X}$, we can define $\mathbf{Y} = \mathbf{X} - \mu_{\mathbf{X}}$ so that

$$\mathbf{C}_{\mathbf{X}} = \mathrm{E}\left[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})'\right] = \mathrm{E}\left[\mathbf{Y}\mathbf{Y}'\right] = \mathbf{R}_{\mathbf{Y}}. \tag{1}$$

It follows that the covariance matrix $\mathbf{C}_{\mathbf{X}}$ is positive semi-definite if and only if the correlation matrix $\mathbf{R}_{\mathbf{Y}}$ is positive semi-definite. Thus, it is sufficient to show that every correlation matrix, whether it is denoted $\mathbf{R}_{\mathbf{Y}}$ or $\mathbf{R}_{\mathbf{X}}$, is positive semi-definite.

To show a correlation matrix $\mathbf{R}_{\mathbf{X}}$ is positive semi-definite, we write

$$\mathbf{a}'\mathbf{R}_{\mathbf{X}}\mathbf{a} = \mathbf{a}'\,\mathrm{E}\left[\mathbf{X}\mathbf{X}'\right]\mathbf{a} = \mathrm{E}\left[\mathbf{a}'\mathbf{X}\mathbf{X}'\mathbf{a}\right] = \mathrm{E}\left[(\mathbf{a}'\mathbf{X})(\mathbf{X}'\mathbf{a})\right] = \mathrm{E}\left[(\mathbf{a}'\mathbf{X})^2\right]. \tag{2}$$

We note that $W = \mathbf{a}'\mathbf{X}$ is a random variable. Since $\mathrm{E}[W^2] \geq 0$ for any random variable W,

$$\mathbf{a}'\mathbf{R}_{\mathbf{X}}\mathbf{a} = \mathrm{E}\left[W^2\right] \geq 0. \tag{3}$$

Problem 8.5.5

(a) $\mathbf{C}$ must be symmetric since

$$\alpha = \beta = \mathrm{E}\left[X_1 X_2\right]. \tag{1}$$

In addition, α must be chosen so that $\mathbf{C}$ is positive semi-definite. Since the characteristic equation is

$$\begin{aligned}\det(\mathbf{C} - \lambda\mathbf{I}) &= (1-\lambda)(4-\lambda) - \alpha^2 \\ &= \lambda^2 - 5\lambda + 4 - \alpha^2 = 0,\end{aligned} \tag{2}$$

the eigenvalues of $\mathbf{C}$ are

$$\lambda_{1,2} = \frac{5 \pm \sqrt{25 - 4(4-\alpha^2)}}{2}. \tag{3}$$

The eigenvalues are non-negative as long as $\alpha^2 \leq 4$, or $|\alpha| \leq 2$. Another way to reach this conclusion is through the requirement that $|\rho_{X_1 X_2}| \leq 1$.

(b) It remains true that $\alpha = \beta$ and $\mathbf{C}$ must be positive semi-definite. For $\mathbf{X}$ to be a Gaussian vector, $\mathbf{C}$ also must be positive definite. For the eigenvalues of $\mathbf{C}$ to be strictly positive, we must have $|\alpha| < 2$.

(c) Since $\mathbf{X}$ is a Gaussian vector, W is a Gaussian random variable. Thus, we need only calculate

$$\mathrm{E}[W] = 2\,\mathrm{E}[X_1] - \mathrm{E}[X_2] = 0, \tag{4}$$

and

$$\begin{aligned}\mathrm{Var}[W] = \mathrm{E}\left[W^2\right] &= \mathrm{E}\left[4X_1^2 - 4X_1X_2 + X_2^2\right] \\ &= 4\,\mathrm{Var}[X_1] - 4\,\mathrm{Cov}[X_1, X_2] + \mathrm{Var}[X_2] \\ &= 4 - 4\alpha + 4 = 4(2-\alpha). \end{aligned} \tag{5}$$

The PDF of W is

$$f_W(w) = \frac{1}{\sqrt{8(2-\alpha)\pi}} e^{-w^2/8(2-\alpha)}. \tag{6}$$

Problem 8.5.13

As given in the problem statement, we define the m-dimensional vector $\mathbf{X}$, the n-dimensional vector $\mathbf{Y}$ and $\mathbf{W} = \begin{bmatrix}\mathbf{X}' \\ \mathbf{Y}'\end{bmatrix}'$. Note that $\mathbf{W}$ has expected value

$$\mu_{\mathbf{W}} = \mathrm{E}[\mathbf{W}] = \mathrm{E}\left[\begin{bmatrix}\mathbf{X} \\ \mathbf{Y}\end{bmatrix}\right] = \begin{bmatrix}\mathrm{E}[\mathbf{X}] \\ \mathrm{E}[\mathbf{Y}]\end{bmatrix} = \begin{bmatrix}\mu_{\mathbf{X}} \\ \mu_{\mathbf{Y}}\end{bmatrix}. \tag{1}$$

The covariance matrix of $\mathbf{W}$ is

$$\begin{aligned}\mathbf{C_W} &= \mathrm{E}\left[(\mathbf{W} - \mu_{\mathbf{W}})(\mathbf{W} - \mu_{\mathbf{W}})'\right] \\ &= \mathrm{E}\left[\begin{bmatrix}\mathbf{X} - \mu_{\mathbf{X}} \\ \mathbf{Y} - \mu_{\mathbf{Y}}\end{bmatrix}\begin{bmatrix}(\mathbf{X} - \mu_{\mathbf{X}})' & (\mathbf{Y} - \mu_{\mathbf{Y}})'\end{bmatrix}\right] \\ &= \begin{bmatrix}\mathrm{E}[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})'] & \mathrm{E}[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{Y} - \mu_{\mathbf{Y}})'] \\ \mathrm{E}[(\mathbf{Y} - \mu_{\mathbf{Y}})(\mathbf{X} - \mu_{\mathbf{X}})'] & \mathrm{E}[(\mathbf{Y} - \mu_{\mathbf{Y}})(\mathbf{Y} - \mu_{\mathbf{Y}})']\end{bmatrix} = \begin{bmatrix}\mathbf{C_X} & \mathbf{C_{XY}} \\ \mathbf{C_{YX}} & \mathbf{C_Y}\end{bmatrix}. \end{aligned} \tag{2}$$

The assumption that $\mathbf{X}$ and $\mathbf{Y}$ are independent implies that

$$\mathbf{C_{XY}} = \mathrm{E}\left[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{Y}' - \mu_{\mathbf{Y}}')\right] = (\mathrm{E}[(\mathbf{X} - \mu_{\mathbf{X}})]\,\mathrm{E}\left[(\mathbf{Y}' - \mu_{\mathbf{Y}}')\right] = \mathbf{0}. \tag{3}$$

This also implies $\mathbf{C_{YX}} = \mathbf{C}_{\mathbf{XY}}' = \mathbf{0}'$. Thus

$$\mathbf{C_W} = \begin{bmatrix}\mathbf{C_X} & \mathbf{0} \\ \mathbf{0}' & \mathbf{C_Y}\end{bmatrix}. \tag{4}$$

Problem 8.5.14

(a) If you are familiar with the Gram-Schmidt procedure, the argument is that applying Gram-Schmidt to the rows of $\mathbf{A}$ yields m orthogonal row vectors. It is then possible to augment those vectors with an additional $n-m$ orothogonal vectors. Those orthogonal vectors would be the rows of $\tilde{\mathbf{A}}$.
An alternate argument is that since $\mathbf{A}$ has rank m the nullspace of $\mathbf{A}$, i.e., the set of all vectors $\mathbf{y}$ such that $\mathbf{Ay} = \mathbf{0}$ has dimension $n-m$. We can choose any $n-m$ linearly independent vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{n-m}$ in the nullspace $\mathbf{A}$. We then define $\tilde{\mathbf{A}}'$ to have columns $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{n-m}$. It follows that $\mathbf{A}\tilde{\mathbf{A}}' = \mathbf{0}$.

(b) To use Theorem 8.11 for the case $m = n$ to show

$$\bar{\mathbf{Y}} = \begin{bmatrix} \mathbf{Y} \\ \hat{\mathbf{Y}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \hat{\mathbf{A}} \end{bmatrix} \mathbf{X}. \tag{1}$$

is a Gaussian random vector requires us to show that

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A} \\ \hat{\mathbf{A}} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1} \end{bmatrix} \tag{2}$$

is a rank n matrix. To prove this fact, we will suppose there exists $\mathbf{w}$ such that $\bar{\mathbf{A}}\mathbf{w} = \mathbf{0}$, and then show that $\mathbf{w}$ is a zero vector. Since $\mathbf{A}$ and $\tilde{\mathbf{A}}$ together have n linearly independent rows, we can write the row vector $\mathbf{w}'$ as a linear combination of the rows of $\mathbf{A}$ and $\tilde{\mathbf{A}}$. That is, for some $\mathbf{v}$ and $\tilde{\mathbf{v}}$,

$$\mathbf{w}' = \mathbf{v}\mathbf{t}'\mathbf{A} + \tilde{\mathbf{v}}'\tilde{\mathbf{A}}. \tag{3}$$

The condition $\bar{\mathbf{A}}\mathbf{w} = \mathbf{0}$ implies

$$\begin{bmatrix} \mathbf{A} \\ \tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1} \end{bmatrix} (\mathbf{A}'\mathbf{v} + \tilde{\mathbf{A}}'\tilde{\mathbf{v}}') = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{4}$$

This implies

$$\mathbf{A}\mathbf{A}'\mathbf{v} + \mathbf{A}\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = \mathbf{0}, \tag{5}$$

$$\tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1}\mathbf{A}\mathbf{v} + \tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1}\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = \mathbf{0}. \tag{6}$$

Since $\mathbf{A}\tilde{\mathbf{A}}' = \mathbf{0}$, it follows that $\mathbf{A}\mathbf{A}'\mathbf{v} = \mathbf{0}$. Since $\mathbf{A}$ is rank m, $\mathbf{A}\mathbf{A}'$ is an $m \times m$ rank m matrix. It follows that $\mathbf{v} = \mathbf{0}$. We can then conclude that

$$\tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1}\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = \mathbf{0}. \tag{7}$$

This would imply that $\tilde{\mathbf{v}}'\tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1}\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = 0$. Since $\mathbf{C}_{\mathbf{X}}^{-1}$ is invertible, this would imply that $\tilde{\mathbf{A}}'\tilde{\mathbf{v}} = \mathbf{0}$. Since the rows of $\tilde{\mathbf{A}}$ are linearly independent, it must be that $\tilde{\mathbf{v}} = \mathbf{0}$. Thus $\bar{\mathbf{A}}$ is full rank and $\bar{\mathbf{Y}}$ is a Gaussian random vector.

(c) We note that By Theorem 8.11, the Gaussian vector $\bar{\mathbf{Y}} = \bar{\mathbf{A}}\mathbf{X}$ has covariance matrix

$$\bar{\mathbf{C}} = \bar{\mathbf{A}}\mathbf{C}_{\mathbf{X}}\bar{\mathbf{A}}'. \tag{8}$$

Since $(\mathbf{C}_{\mathbf{X}}^{-1})' = \mathbf{C}_{\mathbf{X}}^{-1}$,

$$\bar{\mathbf{A}}' = \begin{bmatrix} \mathbf{A}' & (\tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1})' \end{bmatrix} = \begin{bmatrix} \mathbf{A}' & \mathbf{C}_{\mathbf{X}}^{-1}\tilde{\mathbf{A}}' \end{bmatrix}. \tag{9}$$

Applying this result to Equation (8) yields

$$\begin{aligned} \bar{\mathbf{C}} &= \begin{bmatrix} \mathbf{A} \\ \tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1} \end{bmatrix} \mathbf{C}_{\mathbf{X}} \begin{bmatrix} \mathbf{A}' & \mathbf{C}_{\mathbf{X}}^{-1}\tilde{\mathbf{A}}' \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}\mathbf{C}_{\mathbf{X}} \\ \tilde{\mathbf{A}} \end{bmatrix} \begin{bmatrix} \mathbf{A}' & \mathbf{C}_{\mathbf{X}}^{-1}\tilde{\mathbf{A}}' \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}' & \mathbf{A}\tilde{\mathbf{A}}' \\ \tilde{\mathbf{A}}\mathbf{A}' & \tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1}\tilde{\mathbf{A}}' \end{bmatrix}. \end{aligned} \tag{10}$$

Since $\tilde{\mathbf{A}}\mathbf{A}' = \mathbf{0}$,

$$\bar{\mathbf{C}} = \begin{bmatrix} \mathbf{A}\mathbf{C}_{\mathbf{X}}\mathbf{A}' & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{A}}\mathbf{C}_{\mathbf{X}}^{-1}\tilde{\mathbf{A}}' \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{\mathbf{Y}} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\hat{\mathbf{Y}}} \end{bmatrix}. \tag{11}$$

We see that $\bar{\mathbf{C}}$ is block diagonal covariance matrix. From the claim of Problem 8.5.13, we can conclude that $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ are independent Gaussian random vectors.