Solutions:

1. Bochner-Khinchin Theorem (necessary part) (10 points)

\[
\sum_{i,j=1}^{\infty} \phi(\omega_i \omega_j) \lambda_i \tilde{\lambda}_j = \sum_{i,j=1}^{n} E[\exp(i\omega_i X) \exp(-i\omega_j X)] \lambda_i \tilde{\lambda}_j
\]
\[
= \sum_{i,j=1}^{n} E \left[ \lambda_i \exp(i\omega_i X) \{\exp(i\omega_j X) \lambda_j}\right]
\]
\[
= E \left\{ \sum_{i=1}^{n} \lambda_i \exp(i\omega_i X) \right\}^2 \geq 0.
\]

2. Independent Gaussian random variables (10 points)

For simplicity, first we assume \( \mu_X = \mu_Y = 0 \). Then,
\[
\text{cov}(aX - bY, aX + bY) = a^2 \text{cov}(X) - b^2 \text{cov}(Y) = a^2 \sigma_X^2 - b^2 \sigma_Y^2
\]
given that \( X, Y \) are independent Gaussian random variables. Therefore,

(a) \( X + Y, X - Y \) are independent if \( \sigma_X = \sigma_Y \).

(b) \( aX + bY, aX - bY \) are independent for any pair \( (a, b) \in \mathbb{R}^2 \) if \( a^2 \sigma_X^2 = b^2 \sigma_Y^2 \).

3. Gaussian random vector (10 points)

(a) From \( X = AZ \), we have that
\[
\text{cov}(X) = A\text{cov}(Z)A^T.
\]
Since \( Z \) is a Gaussian random vector with zero mean and covariance being the identity matrix, we need to find a matrix \( A \) such that \( \Sigma = AA^T \). Such a decomposition is always possible since \( \Sigma \) is positive definite. One way to obtain a solution is through eigenvalue decomposition. If \( \lambda_1, \lambda_2 \) and \( \lambda_3 \) are the eigenvalues of \( \Sigma \) and \( \xi_1, \xi_2 \) and \( \xi_3 \) are the corresponding unit right-eigenvectors, then
\[
A = [\xi_1 \xi_2 \xi_3] \begin{bmatrix}
\sqrt{\lambda_1} & 0 & 0 \\
0 & \sqrt{\lambda_2} & 0 \\
0 & 0 & \sqrt{\lambda_3}
\end{bmatrix}
\]
is a solution.

For the problem at hand, we can easily show that \( \lambda_1 = 2, \lambda_2 = 2 + \sqrt{2}, \lambda_3 = 2 - \sqrt{2} \), and that \( \xi_1 = [\frac{1}{\sqrt{2}} \ - \frac{1}{\sqrt{2}}]^T, \xi_2 = [\frac{1}{\sqrt{2}} \ \frac{1}{\sqrt{2}}]^T \), and \( \xi_3 = [\frac{1}{\sqrt{2}} \ - \frac{1}{\sqrt{2}}]^T \). The matrix \( A \) is obtained as
\[
A = \begin{bmatrix}
0 & \frac{\sqrt{2} + \sqrt{2}}{\sqrt{2}} & -\frac{\sqrt{2} \ - \sqrt{2}}{\sqrt{2}} \\
1 & \frac{\sqrt{2} + \sqrt{2}}{\sqrt{2}} & \frac{\sqrt{2} \ - \sqrt{2}}{\sqrt{2}} \\
-1 & \frac{\sqrt{2} + \sqrt{2}}{\sqrt{2}} & \frac{\sqrt{2} \ - \sqrt{2}}{\sqrt{2}}
\end{bmatrix}
\] \hspace{1cm} (2)

Another solution for \( A \) can be obtained by Cholesky decomposition.
(b) Since $X_1, X_2, X_3$ are zero mean Gaussian random vector, we know that

$$ E[X_1|X_2, X_3] = aX_2 + bX_3. $$

Also by orthogonal principle,

$$ E[X_i(X_1 - aX_2 - bX_3)] = 0; \ i = 2, 3. $$

i.e., $1 - 2a = 0$ and $1 - 2b = 0$.

$$ \Rightarrow a = b = \frac{1}{2}. \quad (3) $$

4. **Gaussian related distributions** (20 points)

Let $X_1, X_2, \ldots, X_n$ be independent variables, $X_i$ being $N(\mu_i, 1)$, and let $Y = (R^2 =) X_1^2 + X_2^2 + \cdots + X_n^2$.

(a) If $X$ is $N(\mu_0, 1)$, then the moment generating function of $X^2$ is

$$ M_{X^2}(s) = E(e^{sx^2}) = \int_{-\infty}^{\infty} e^{sx^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu_0)^2} dx = \frac{1}{\sqrt{1-2s}} \exp \left( \frac{\mu_0^2}{1-2s} \right), $$

if $s < \frac{1}{2}$, by completing the square in the exponent. It follows that

$$ M_Y(s) = \prod_{j=1}^{n} \left\{ \frac{1}{\sqrt{1-2s}} \exp \left( \frac{\mu_j^2}{1-2s} \right) \right\} = \frac{1}{(1-2s)^{n/2}} \exp \left( \frac{\theta s}{1-2s} \right), $$

which shows that the characteristic function of $Y$ is

$$ \phi_Y(\omega) = \frac{1}{(1-2i\omega)^{n/2}} \exp \left( \frac{i\omega\theta}{1-2i\omega} \right) \quad (4) $$

where $\theta = \mu_1^2 + \cdots + \mu_n^2$.

(b) The transformation $X_1 = R \cos \Theta, X_2 = R \sin \Theta$ has Jacobian $J = r$, so that

$$ f_{R,\Theta}(r, \theta) = \frac{1}{2\pi} r e^{-\frac{1}{2}r^2}, \ r > 0, 0 \leq \theta < 2\pi. $$

Therefore, $R$ and $\Theta$ are independent, $\Theta$ being uniform on $[0, 2\pi)$, and $R^2$ having distribution function

$$ P(R^2 \leq a) = \int_0^{\sqrt{a}} r e^{-\frac{1}{2}r^2} dr = 1 - e^{-\frac{1}{2}a}; $$

this is the exponential distribution with parameter $\frac{1}{2}$. The density function of $R$ is

$$ f_R(r) = re^{-\frac{1}{2}r^2}, \ r > 0. $$
(c) By symmetry,

\[ E \left( \frac{X_1^2}{R^2} \right) = \frac{1}{2} E \left( \frac{X_1^2 + X_2^2}{R^2} \right) = \frac{1}{2}. \]

In the first octant, i.e., in \( \{ (x_1, x_2) : 0 \leq x_2 \leq x_1 \} \), we have

\[ \min\{x_1, x_2\} = x_2, \max\{x_1, x_2\} = x_1. \]

The joint density \( f_{X_1, X_2} \) is invariant under rotations, and hence the expectation in question is

\[
E \left[ \frac{\min\{|X_1|, |X_2|\}}{\max\{|X_1|, |X_2|\}} \right] = 8 \int_0^{\pi/4} \int_0^{\infty} \frac{x_2}{2} f_{X_1, X_2}(x_1, x_2) dx_1 dx_2 = \frac{2}{\pi} \log 2.
\]

5. **Asymptotic property of characteristic function** (10 points)

Let \( \epsilon > 0 \) and using Riemann-Lebesgue lemma, we choose \( g_\epsilon(t) \) to be a (finite) step function such that

\[ \int_{-\infty}^{\infty} |f_X(t) - g_\epsilon(t)| dt < \epsilon. \]

Then, the following inequality holds

\[ \left| \int_{-\infty}^{\infty} e^{i\omega x} (f_X(x) - g_\epsilon(x)) dx \right| \leq \int_{-\infty}^{\infty} \left| e^{i\omega x} \right| |f_X(x) - g_\epsilon(x)| dx = \int_{-\infty}^{\infty} |f_X(x) - g_\epsilon(x)| dx \leq \epsilon, \]

i.e.

\[ \left| \phi_X(\omega) - \int_{-\infty}^{\infty} e^{i\omega x} g_\epsilon(x) dx \right| < \epsilon. \quad (5) \]

Note that \( g_\epsilon(x) \) can be written as

\[ g_\epsilon(x) = \sum_{i=1}^{\infty} \alpha_i I_{[a_i, b_i]} \]

for some \( \alpha_i, a_i, b_i \)'s. Apparently,

\[ \lim_{\omega \to \infty} \left| \int_{-\infty}^{\infty} e^{i\omega x} I_{[a_i, b_i]} dx \right| \leq \lim_{\omega \to \infty} \frac{1}{\omega} \left| e^{i\omega b_i} - e^{i\omega a_i} \right| = 0. \]

From (5), it follows

\[ \lim_{\omega \to \infty} \phi_X(\omega) = \lim_{\omega \to \infty} \int_{-\infty}^{\infty} e^{i\omega x} g_\epsilon(x) dx = 0. \]

6. **Symmetric matrices** (20 points)

(a) 5.11 in the textbook, page 300.
i. \( A^T K A = I \Rightarrow AA^T K A = A \). Also,
\[
A^T K_1 A = \Lambda^{(1)} \Rightarrow AA^T K_1 A = A A A^T K A = (AA^T K A) \Lambda^{(1)}. 
\]
It follows that \( AA^T (K_1 A - K A A) = 0 \). Therefore, \( K_1 A = K A A \), or equivalently,
\[
K^{-1} K_1 A = AA \Lambda^{(1)}. 
\]
ii. \( A^T K A = A^T (a_1 K_1 + a_2 K_2) A = a_1 A^T K_1 A + a_2 A^T K_2 A = I \). Then,
\[
A^T K_2 A = \frac{1}{a_2} \begin{pmatrix}
1 - a_1 \lambda_1^{(1)} \\
\vdots \\
\vdots \\
1 - a_1 \lambda_n^{(1)}
\end{pmatrix}
\begin{pmatrix}
\lambda_1^{(2)} \\
\vdots \\
\vdots \\
\lambda_n^{(2)}
\end{pmatrix}
\]
where \( \lambda_i^{(2)} = \frac{1}{a_2}(1 - a_1 \lambda_i^{(1)}) \), \( i = 1, \ldots, n \).

iii. For any eigenvector \( v_i \) of \( A^T K_1 A \), we have \( A^T K_1 A v_i = \lambda_i^{(1)} v_i \). Then,
\[
\frac{1}{a_1} A^T (K - a_2 K_2) A v_i = \frac{1}{a_1} A^T K A v_i - \frac{a_2}{a_1} A^T K_2 A v_i = \frac{1}{a_1} v_i - \frac{a_2}{a_1} A^T K_2 A v_i = \lambda_i^{(1)} v_i
\]
Therefore, \( A^T K_2 A v_i = \frac{1}{a_2} (1 - a_1 \lambda_i^{(1)}) v_i = \lambda_i^{(2)} v_i \), and \( v_i \) is also an eigenvector of \( A^T K_2 A \), with corresponding eigenvalue \( \lambda_i^{(2)} = \frac{1}{a_2} (1 - a_1 \lambda_i^{(1)}) \).

iv. From previous derivations, we already have \( \lambda_i^{(2)} = \frac{1}{a_2} (1 - a_1 \lambda_i^{(1)}) \).

(b) 5.12 in the textbook, page 300.

i. \( \frac{1}{m} W W^T = \frac{1}{m} (X_1 \cdots X_n) \begin{pmatrix} X_1^T \\ \vdots \\ X_n^T \end{pmatrix} = \frac{1}{m} \sum_{i=1}^{m} X_i X_i^T = S \)

ii. \( \text{rank}(S) \leq \min\{n, m\} = m \). The maximum rank of \( S \) is \( m \).

iii. Note that the size of \( S' \) is \( m \times m \). We have the following:
\[
S' \Phi = \Phi A \Rightarrow \frac{1}{m} W^T W \Phi = \Phi A \\
\Rightarrow W (\frac{1}{m} W^T W \Phi) = W \Phi A \\
\Rightarrow (\frac{1}{m} W W^T) (W \Phi) = (W \Phi) A \\
\Rightarrow S (W \Phi) = (W \Phi) A
\]
So, \( S \) and \( S' \) have the same (non-zero) eigenvalues. For any eigenvector \( v_i \) of \( S' \), the matrix \( S \) has a corresponding eigenvector \( u_i = W v_i \).

iv. The advantage is that \( S' \) has smaller size (\( m \times m \) instead of \( n \times n \)), thus taking fewer computations.
7. Convergence of random variables (10 points)

(a) First, if \( p \geq 1 \), then using Markov inequality, we have

\[
P(|X_n - X| > \epsilon) \leq \frac{E[|X_n - X|]}{\epsilon} \leq \frac{(E[|X_n - X|^p])^{1/p}}{\epsilon} \to 0
\]

where the second inequality comes from Jensen’s inequality (remember \( f(\cdot)^p, p \geq 1 \), is a convex function). While for \( 0 < p < 1 \), we have

\[
|X_n - X| < \epsilon < 1 \Rightarrow E[|X_n - X|] \leq E[|X_n - X|^p].
\]

It follows that \( E[|X_n - X|] \to 0 \) and \( X_n \) converges to \( X \) in probability.

(b) If \( \lim_{n \to \infty} X_n \to X(\cdot) \) and \( \lim_{n \to \infty} X_n \to Y(\cdot) \), then

\[
P(|X - Y| > \epsilon) \leq P(|X_n - X| + |X_n - Y| > \epsilon) \leq P(|X_n - X| > \epsilon) + P(|X_n - Y| > \epsilon) \to 0.
\]

Therefore, \( P(X \neq Y) = 0 \), i.e. \( X = Y \) (a.s.);

(c) Let \( A = \{\omega : \lim_{n \to \infty} |X_n(\omega) - X(\omega)|^2 = 0\} \), and \( B = \{\omega : \lim_{n \to \infty} |X_n^2(\omega) - X^2(\omega)| = 0\} \).

Then, \( \forall \omega \in A \),

\[
\lim_{n \to \infty} |X_n(\omega) - X(\omega)|^2 = 0 \Rightarrow \lim_{n \to \infty} |X_n(\omega) - X(\omega)| = 0 \Rightarrow \lim_{n \to \infty} X_n^2(\omega) = X^2(\omega)
\]

it follows that \( A \subseteq B \) and \( P(B) \geq P(A) = 1 \). Hence, \( P(B) = 1 \), and \( X_n^2 \to X^2 \) (a.s.).