S&DS 351: Stochastic Processes - Homework 7

Bryan SebaRaj

Professor Ilias Zadik

April 9, 2025

Problem 1. Let X and Y be two jointly (non-degenerate) Gaussian random variables with mean 0 and covariance Σ (so $(X,Y) \sim \mathcal{N}(0,\Sigma)$). Thus, $\sigma_X^2 = \Sigma_{11} > 0$, $\sigma_Y^2 = \Sigma_{22} > 0$, and $\mathbb{E}[XY] = \Sigma_{12}$.

In this problem we will prove that we can always write for some real number a, Y = aX + V where X and V are independent Gaussian random variables — a very handy formula in applications.

(a) (10 points) Here we guess the correct value of a. Assuming that X and V are independent and Y = aX + V, prove that it must hold $\Sigma_{12} = a\Sigma_{11}$, or $a = \Sigma_{12}(\Sigma_{11})^{-1}$.

Compute the covariance between X and Y,

$$\mathbb{E}[XY] = \mathbb{E}[X(aX + V)] = \mathbb{E}[aX^2] + \mathbb{E}[XV]$$

Since X and V are independent and $\mathbb{E}[X] = 0$,

$$\mathbb{E}[XY] = a\mathbb{E}[X^2] + \mathbb{E}[X]\mathbb{E}[V] = a\sigma_X^2 + 0 \cdot \mathbb{E}[V] = a\Sigma_{11}$$

Given that $\mathbb{E}[XY] = \Sigma_{12}$,

$$\Sigma_{12} = a\Sigma_{11}$$

(b) (10 points) Here we guess the correct values of the mean and variance of V. Assuming that X and V are independent and Y = aX + V, compute the necessary values of mean and variance of V in terms of $\Sigma_{11}, \Sigma_{22}, \Sigma_{12}$.

Since Y = aX + V and $\mathbb{E}[Y] = 0$,

$$\mathbb{E}[Y] = \mathbb{E}[aX + V]$$

$$0 = a\mathbb{E}[X] + \mathbb{E}[V]$$

Since $\mathbb{E}[X] = 0$,

$$\mathbb{E}[V] = 0$$

In order to determine the variance of V,

$$Var(Y) = Var(aX + V)$$

Since X and V are independent,

$$Var(Y) = a^2 Var(X) + Var(V)$$

$$\Sigma_{22} = a^2 \Sigma_{11} + \text{Var}(V)$$

$$Var(V) = \Sigma_{22} - a^2 \Sigma_{11}$$

$$Var(V) = \Sigma_{22} - \left(\frac{\Sigma_{12}}{\Sigma_{11}}\right)^2 \Sigma_{11} = \Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}$$

Therefore, V has mean $\mathbb{E}[V] = 0$ and variance $\text{Var}(V) = \Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}$.

(c) (10 points) Assume that X is any random variable and V is an independent random variable from X which is Gaussian with mean and variance from part (b) as a function of $\Sigma_{11}, \Sigma_{22}, \Sigma_{12}$. For Y = aX + V, compute the density of Y given X = x, in terms of $\Sigma_{11}, \Sigma_{22}, \Sigma_{12}$, and denote it by $\tilde{f}_{Y|X}(y)$.

Since V is independent of X and follows a Gaussian distribution with mean 0 and variance $\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}$,

$$Y|X = x \sim \mathcal{N}\left(ax, \Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)$$

Calculating the conditional density function,

$$\tilde{f}_{Y|X}(y) = \frac{1}{\sqrt{2\pi \left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}} \exp\left(-\frac{(y - ax)^2}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}\right)$$

Substituting $a = \frac{\Sigma_{12}}{\Sigma_{11}}$,

$$\tilde{f}_{Y|X}(y) = \frac{1}{\sqrt{2\pi \left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}} \exp\left(-\frac{(y - \frac{\Sigma_{12}}{\Sigma_{11}}x)^2}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}\right)$$

(d) (10 points) Recall the pdf of the joint Gaussian distribution of X and Y that we presented in class f_{XY} . Verify that for the choice of V from (b) and the resulting density $\tilde{f}_{Y|X}(y)$ from part (c) it holds

$$f_{XY}(x,y) = f_X(x)\tilde{f}_{Y|X}(y)$$

for all $x, y \in \mathbb{R}$. Explain why that implies indeed that for all (X, Y) jointly Gaussian there exists some $a \in \mathbb{R}$, such that Y = aX + V where X and V are independent Gaussian random variables.

For jointly Gaussian random variables X and Y with mean 0 and covariance matrix Σ , the joint probability density function is:

$$f_{XY}(x,y) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}[x\ y]\Sigma^{-1}\begin{bmatrix}x\\y\end{bmatrix}\right)$$

Given that,

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}$$

The determinant can be calculated as $det(\Sigma) = \Sigma_{11}\Sigma_{22} - \Sigma_{12}^2$. The inverse of Σ can be given as

$$\Sigma^{-1} = \frac{1}{\det(\Sigma)} \begin{bmatrix} \Sigma_{22} & -\Sigma_{12} \\ -\Sigma_{12} & \Sigma_{11} \end{bmatrix}$$

Computing the exponent in the joint PDF,

$$[x \ y] \Sigma^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{\det(\Sigma)} [x \ y] \begin{bmatrix} \Sigma_{22} & -\Sigma_{12} \\ -\Sigma_{12} & \Sigma_{11} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$= \frac{1}{\det(\Sigma)} [x \Sigma_{22} - y \Sigma_{12} - x \Sigma_{12} + y \Sigma_{11}] \begin{bmatrix} x \\ y \end{bmatrix}$$
$$= \frac{1}{\det(\Sigma)} (x^2 \Sigma_{22} - xy \Sigma_{12} - xy \Sigma_{12} + y^2 \Sigma_{11})$$
$$= \frac{1}{\det(\Sigma)} (x^2 \Sigma_{22} - 2xy \Sigma_{12} + y^2 \Sigma_{11})$$

Therefore, the joint PDF is

$$f_{XY}(x,y) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2} \frac{x^2 \Sigma_{22} - 2xy \Sigma_{12} + y^2 \Sigma_{11}}{\det(\Sigma)}\right)$$

Note that the marginal denisty of X is

$$f_X(x) = \frac{1}{\sqrt{2\pi\Sigma_{11}}} \exp\left(-\frac{x^2}{2\Sigma_{11}}\right)$$

Computing $f_X(x) \cdot \tilde{f}_{Y|X}(y)$,

$$f_X(x) \cdot \tilde{f}_{Y|X}(y) = \frac{1}{\sqrt{2\pi\Sigma_{11}}} \exp\left(-\frac{x^2}{2\Sigma_{11}}\right) \cdot \frac{1}{\sqrt{2\pi\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}} \exp\left(-\frac{(y - \frac{\Sigma_{12}}{\Sigma_{11}}x)^2}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}\right)$$

$$= \frac{1}{2\pi\sqrt{\Sigma_{11}\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}} \exp\left(-\frac{x^2}{2\Sigma_{11}} - \frac{(y - \frac{\Sigma_{12}}{\Sigma_{11}}x)^2}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}\right)$$

Simplying the denominator within the square root,

$$\Sigma_{11} \left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}} \right) = \Sigma_{11} \Sigma_{22} - \Sigma_{12}^2 = \det(\Sigma)$$

Therefore,

$$f_X(x) \cdot \tilde{f}_{Y|X}(y) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left(-\frac{x^2}{2\Sigma_{11}} - \frac{(y - \frac{\Sigma_{12}}{\Sigma_{11}}x)^2}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}\right)$$

Note that,

$$\frac{(y - \frac{\Sigma_{12}}{\Sigma_{11}}x)^2}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)} = \frac{y^2 - 2y\frac{\Sigma_{12}}{\Sigma_{11}}x + \frac{\Sigma_{12}^2}{\Sigma_{11}^2}x^2}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)}$$

Combining the exponents,

$$-\frac{x^2}{2\Sigma_{11}} - \frac{(y - \frac{\Sigma_{12}}{\Sigma_{11}}x)^2}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)} = -\frac{x^2}{2\Sigma_{11}} - \frac{1}{2\left(\Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)} \left(y^2 - 2y\frac{\Sigma_{12}}{\Sigma_{11}}x + \frac{\Sigma_{12}^2}{\Sigma_{11}^2}x^2\right)$$

Combining like terms, the exponent becomes

$$-\frac{1}{2} \frac{x^2 \Sigma_{22} - 2xy \Sigma_{12} + y^2 \Sigma_{11}}{\det(\Sigma)}$$

Note that this is exactly the exponent in the joint PDF $f_{XY}(x,y)$. As such,

$$f_{XY}(x,y) = f_X(x) \cdot \tilde{f}_{Y|X}(y)$$

This verifies that the construction of Y = aX + V when $a = \frac{\Sigma_{12}}{\Sigma_{11}}$ and $V \sim \mathcal{N}\left(0, \Sigma_{22} - \frac{\Sigma_{12}^2}{\Sigma_{11}}\right)$ independent of X indeed yields the correct joint distribution of (X,Y). The fact that $f_{XY}(x,y) = f_X(x) \cdot \tilde{f}_{Y|X}(y)$ implies that $\tilde{f}_{Y|X}(y)$ is the true conditional density of Y given X = x. This confirms that for any jointly Gaussian random variables X and Y, Y = aX + V where Y is a Gaussian random variable independent of X.

Problem 2. (20 points)

Prove that the real-valued random variables X_1, \ldots, X_n follow a joint distribution which is Gaussian if and only if for all $a_1, \ldots, a_n \in \mathbb{R}$, $a_1 X_1 + \cdots + a_n X_n$ follows a Gaussian distribution.

Note: You can use that an m-dimensional distribution is a Gaussian on m dimensions if and only if for some $\mu \in \mathbb{R}^m$, $\Sigma \in \mathbb{R}^{m \times m}$ the MGF of the distribution equals $\exp\left(a^T \mu + a^T \Sigma a/2\right)$ for all $a \in \mathbb{R}^m$.

First, proving the forward direction. Suppose that (X_1, \ldots, X_n) is jointly Gaussian. Then by definition, there exists $\mu \in \mathbb{R}^n$ and a symmetric positive semidefinite matrix $\Sigma \in \mathbb{R}^{n \times n}$ s.t. the MGF of $X = (X_1, \ldots, X_n)^T$ satisfies

$$M_X(a) = \mathbb{E}\left(e^{a^T X}\right) = \exp\left(a^T \mu + \frac{1}{2}a^T \Sigma a\right) \quad \forall a \in \mathbb{R}^n$$

For any fixed $a = (a_1, \dots, a_n)^T \in \mathbb{R}^n$, consider the random variable,

$$Y = a^T X = a_1 X_1 + \dots + a_n X_n.$$

Note that its MGF is given by

$$\mathbb{E}\left(e^{\lambda Y}\right) = \mathbb{E}\left(e^{\lambda a^T X}\right) = \exp\left(\lambda a^T \mu + \frac{1}{2}\lambda^2 a^T \Sigma a\right), \quad \forall \lambda \in \mathbb{R}.$$

Since this is the MGF of a one-dimensional Gaussian distribution, it follows that Y is Gaussian.

Thus the forward direction is proven. Now for the reverse.

Assume that for every $a \in \mathbb{R}^n$, the random variable $Y = a^T X$ is Gaussian. Then for each fixed a, there exist $m(a) \in \mathbb{R}$ and $\sigma^2(a) \geq 0$ s.t.

$$\mathbb{E}\left(e^{\lambda a^T X}\right) = \exp\left(\lambda m(a) + \frac{1}{2}\lambda^2 \sigma^2(a)\right) \quad \forall \lambda \in \mathbb{R}$$

In particular, when $\lambda = 1$,

$$\mathbb{E}\left(e^{a^TX}\right) = \exp\left(m(a) + \frac{1}{2}\sigma^2(a)\right)$$

See that the mapping $a \mapsto \mathbb{E}\left(e^{a^TX}\right)$ is the MGF of the random vector X. By the uniqueness theorem for MGFs, there exist $\mu \in \mathbb{R}^n$ and a symmetric, nonnegative, definite matrix $\Sigma \in \mathbb{R}^{n \times n}$ s.t.

$$m(a) = a^T \mu$$
 and $\sigma^2(a) = a^T \Sigma a$, $\forall a \in \mathbb{R}^n$

Therefore,

$$\mathbb{E}\left(e^{a^TX}\right) = \exp\left(a^T\mu + \frac{1}{2}a^T\Sigma a\right) \quad \forall a \in \mathbb{R}^n$$

Note that this is equivalent to the joint distribution of $X = (X_1, \dots, X_n)$ is Gaussian. Thus, X_1, \dots, X_n are jointly Gaussian if and only if every linear combination $a_1X_1 + \dots + a_nX_n$ is Gaussian.

5.3 (10 points)

For 0 < a < b, calculate the conditional probability $P\{W_b > 0 \mid W_a > 0\}$.

Since $\{W_t\}_{t\geq 0}$ is a standard Brownian motion, for any 0 < a < b,

$$W_b = W_a + (W_b - W_a),$$

where $W_a \sim N(0, a)$ and $W_b - W_a \sim N(0, b - a)$, with W_a and $W_b - W_a$ independent. Conditioning on $W_a = x > 0$ given $W_a = x > 0$,

$$P(W_b > 0 \mid W_a = x) = P\{x + (W_b - W_a) > 0\} = P\{W_b - W_a > -x\}$$

Let Φ denote the standard normal cumulative distribution function. Since $W_b - W_a \sim N(0, b - a)$,

$$P(W_b > 0 \mid W_a = x) = \Phi\left(\frac{x}{\sqrt{b-a}}\right)$$

Note that the unconditional distribution of W_a is

$$W_a \sim N(0, a)$$
 with density $f_{W_a}(x) = \frac{1}{\sqrt{2\pi a}} \exp\left(-\frac{x^2}{2a}\right), \quad x \in \mathbb{R}$

Since $P(W_a > 0) = \frac{1}{2}$ by symmetry, the conditional density of W_a given $W_a > 0$ is given by

$$f_{W_a|W_a>0}(x) = \frac{f_{W_a}(x)}{P(W_a>0)} = \frac{2}{\sqrt{2\pi a}} \exp\left(-\frac{x^2}{2a}\right), \quad x>0$$

Therefore,

$$P(W_b > 0 \mid W_a > 0) = \int_0^\infty \Phi\left(\frac{x}{\sqrt{b-a}}\right) f_{W_a \mid W_a > 0}(x) \, dx = \int_0^\infty \Phi\left(\frac{x}{\sqrt{b-a}}\right) \frac{2}{\sqrt{2\pi a}} \exp\left(-\frac{x^2}{2a}\right) dx$$

Since (W_a, W_b) is bivariate normal with mean vector **0** and covariance matrix

$$\Sigma = \begin{pmatrix} a & a \\ a & b \end{pmatrix},$$

the correlation coefficient is given by

$$\rho = \frac{\operatorname{Cov}(W_a, W_b)}{\sqrt{\operatorname{Var}(W_a)\operatorname{Var}(W_b)}} = \frac{a}{\sqrt{a\,b}} = \sqrt{\frac{a}{b}}$$

Note that for bivariate normal random variables,

$$P(W_a > 0, W_b > 0) = \frac{1}{4} + \frac{1}{2\pi}\arcsin(\rho)$$

Substituting $\rho = \sqrt{\frac{a}{b}}$,

$$P(W_a > 0, W_b > 0) = \frac{1}{4} + \frac{1}{2\pi} \arcsin\left(\sqrt{\frac{a}{b}}\right)$$

Since $P(W_b > 0 \mid W_a > 0) = \frac{P(W_a > 0, W_b > 0)}{P(W_a > 0)}$ and $P(W_a > 0) = \frac{1}{2}$,

$$P(W_b > 0 \mid W_a > 0) = \frac{\frac{1}{4} + \frac{1}{2\pi}\arcsin\left(\sqrt{\frac{a}{b}}\right)}{1/2} = \frac{1}{2} + \frac{1}{\pi}\arcsin\left(\sqrt{\frac{a}{b}}\right)$$

Therefore,

$$P(W_b > 0 \mid W_a > 0) = \frac{1}{2} + \frac{1}{\pi}\arcsin\left(\sqrt{\frac{a}{b}}\right)$$

5.4 (15 points)

Prove: Suppose that W is a standard Brownian motion, and let c > 0. Then the process X defined by $X(t) = c^{-1/2}W(ct)$ is also a standard Brownian motion.

Since W(0) = 0 w.p. 1,

$$X(0) = c^{-1/2}W(c \cdot 0) = 0$$

Note that the sample paths of W are continuous. Since the mapping $t \mapsto ct$ is continuous and scaling by $c^{-1/2}$ is a constant multiplication, the process X(t) has continuous sample paths. Let $0 \le s < t$. Considering the increment,

$$X(t) - X(s) = c^{-1/2} (W(ct) - W(cs))$$

Since W(ct) - W(cs) is normally distributed with mean 0 and variance

$$Var(W(ct) - W(cs)) = ct - cs = c(t - s),$$

it follows that

$$X(t) - X(s) \sim N(0, c^{-1} \cdot c(t - s)) = N(0, t - s)$$

Thus, X(t) - X(s) is normally distributed with mean 0 and variance t - s, as required for a standard Brownian motion.

For any partition $0 = t_0 < t_1 < \cdots < t_n$, see that

$$X(t_k) - X(t_{k-1}) = c^{-1/2} \Big(W(ct_k) - W(ct_{k-1}) \Big)$$

Since W has independent increments, $\{W(ct_k) - W(ct_{k-1})\}_{k=1}^n$ are independent. Also note that multiplying by the constant $c^{-1/2}$ preserves independence. Therefore, the increments of X over disjoint intervals are independent.

For $0 \le s \le t$, note that

$$Cov(X(s), X(t)) = Cov(c^{-1/2}W(cs), c^{-1/2}W(ct)) = c^{-1}Cov(W(cs), W(ct))$$

Since for standard Brownian motion, Cov(W(cs), W(ct)) = cs since $s \le t$,

$$Cov(X(s), X(t)) = c^{-1}(cs) = s$$

Thus, X is a standard Brownian motion.

5.6 (15 points)

Prove: Suppose that W is a standard Brownian motion, and let c > 0. Define X(t) = W(c+t) - W(c). Then $\{X(t) : t \ge 0\}$ is a standard Brownian motion that is independent of $\{W(t) : 0 \le t \le c\}$.

Since X(0) = 0 w.p. 1,

$$X(0) = W(c+0) - W(c) = 0$$

Since W has continuous sample paths by definition of Brownian motion, and X(t) = W(c+t) - W(c) is a composition and difference of continuous functions, X also has continuous sample paths. For any $0 \le s < t$ and $0 \le u < v$, consider the increments X(t) - X(s) and X(v) - X(u),

$$X(t) - X(s) = [W(c+t) - W(c)] - [W(c+s) - W(c)] = W(c+t) - W(c+s)$$

$$X(v) - X(u) = [W(c+v) - W(c)] - [W(c+u) - W(c)] = W(c+v) - W(c+u)$$

Note that if the intervals [c+s, c+t] and [c+u, c+v] are disjoint, then the increments X(t) - X(s) and X(v) - X(u) are independent by the independent increments property of the original Brownian motion W.

For stationarity, for any h > 0 and $t \ge 0$, see that

$$X(t+h) - X(t) = W(c+t+h) - W(c) - [W(c+t) - W(c)]$$

= W(c+t+h) - W(c+t)

Since this increment only depends on the time difference h and not starting time t, stationarity is established.

Since W is a Brownian motion, $W(c+t) - W(c) \sim \mathcal{N}(0, (c+t) - c) = \mathcal{N}(0, t)$. Therefore, $X(t) \sim \mathcal{N}(0, t)$ for each t > 0.

Since all properties of standard Brownian motion are satisfied, $\{X(t): t \geq 0\}$ is a standard Brownian motion. Now we need to prove independence, by showing that for any finite collection of times $\{t_1, t_2, \dots, t_n\}$

with $t_i \geq 0$ and $\{s_1, s_2, \ldots, s_m\}$ with $0 \leq s_j \leq c$, the random vectors $(X(t_1), X(t_2), \ldots, X(t_n))$ and $(W(s_1), W(s_2), \ldots, W(s_m))$ are independent.

Suppose $t_0 = 0$ and $s_0 = 0$. Then,

$$X(t_i) = W(c + t_i) - W(c)$$
 for $i = 1, 2, ..., n$
 $W(s_j) = W(s_j) - W(0)$ for $j = 1, 2, ..., m$

Consider the augmented vector of increments,

$$(W(s_1) - W(s_0), W(s_2) - W(s_1), \dots, W(s_m) - W(s_{m-1}), W(c) - W(s_m)$$

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1}))$$

$$(W(s_1) - W(0), W(s_2) - W(s_1), \dots, W(s_m) - W(s_{m-1}), W(c) - W(s_m),$$

$$W(c + t_1) - W(c), W(c + t_2) - W(c + t_1), \dots, W(c + t_n) - W(c + t_{n-1}))$$

By the independent increments property of Brownian motion W, these increments are independent as they are increments over disjoint time intervals.

Since $(W(s_1), W(s_2), \dots, W(s_m))$ can be written as a linear transformation of the first m increments,

$$W(s_1) = W(s_1) - W(0)$$

$$W(s_2) = (W(s_1) - W(0)) + (W(s_2) - W(s_1))$$

Similarly, $(X(t_1), X(t_2), \dots, X(t_n))$ can be written as a linear transformation of the increments involving X,

$$X(t_1) = X(t_1) - X(0) = W(c + t_1) - W(c)$$

$$X(t_2) = X(t_1) + (X(t_2) - X(t_1)) = (W(c + t_1) - W(c)) + (W(c + t_2) - W(c + t_1))$$

Since linear transformations of independent random variables are independent, it follows that the vectors $(W(s_1), W(s_2), \dots, W(s_m))$ and $(X(t_1), X(t_2), \dots, X(t_n))$ are independent. Thus, $\{X(t): t \geq 0\}$ is independent of $\{W(t): 0 \leq t \leq c\}$.