Probability Lecture Notes

Dec 17

Gaussian Random Vectors

Let $(\Omega, \mathcal{F}, P)$ be a Probability Space.

Definition 1. *A random vector $X = (X_1, \ldots, X_m)$ has a **Standard Gaussian Distribution** if its components $X_1, \ldots, X_m$ are jointly i.i.d. random variables with standard normal distribution $\mathcal{N}(0,1)$.*

Definition 2. *A random vector $Y = (Y_1, \ldots, Y_d) : \Omega \to \mathbb{R}^d$ is called a **Gaussian Random Vector** if it is obtained from a vector with standard Gaussian distribution via an affine transformation. That is, there exists a matrix $A \in Mat(d \times m)$ and a vector $b \in \mathbb{R}^d$ such that:*

$$Y^T = AX^T + b^T$$

where $X = (X_1, \ldots, X_m)$ is a standard Gaussian vector.

Note: We treat vectors as row vectors by convention.

Observation 1. *A Gaussian Random Vector need not be independent.*

Recall: For $X, Y : \Omega \to \mathbb{R}$ (real-valued RVs):

$$\text{cov}(X,Y) = E[(X - E(X))(Y - E(Y))] = E(XY) - E(X)E(Y)$$

For vector $Z = (Z_1, \ldots, Z_n)^T : \Omega \to \mathbb{R}^n$, the expectation is the vector $E(Z) = (EZ_1, \ldots, EZ_n)^T \in \mathbb{R}^n$.
The **Covariance Matrix** is defined as:

$$\text{cov}(Z) = \text{cov}(Z,Z) = E[(Z - EZ)^T(Z - EZ)] \in \text{Mat}(n \times n)$$

Lemma 1. *If Y is a Gaussian Random Vector such that $Y^T = AX^T + b^T$, then:*

1. $E(Y) = b$
2. $cov(Y) = AA^T$

Orthogonal Transformations

If $A \in O(d)$ (orthogonal matrices), then $AA^T = A^T A = I$. Orthogonal matrices preserve distances and norms. If X has a standard Gaussian distribution, then AX^T also has a standard Gaussian distribution.

Proof Sketch. Since X is standard Gaussian, its density is μ_X.

$$\mu_X(E) = P(X = (X_1, \dots, X_d) \in E)$$

For a rectangle $E = E_1 \times \cdots \times E_d$, by independence:

$$\mu_X(E) = \mu_{X_1}(E_1) \cdots \mu_{X_d}(E_d) = \int_E F(x)dx$$

where the density is

$$F(x) = \prod_{i=1}^{d} \frac{e^{-x_i^2/2}}{\sqrt{2\pi}} = \frac{e^{-\|x\|^2/2}}{(2\pi)^{d/2}}$$

Then $\mu_{AX}(E) = P(AX \in E) = P(X \in A^{-1}E) = \int_{A^{-1}E} F(x)dx$. Using the change of variables $y = Ax \implies x = A^{-1}y$:

$$\int_E F(A^{-1}y)|\det A^{-1}|dy$$

Since $A \in O(d)$, $|\det A| = |\det A^{-1}| = 1$ and $\|A^{-1}y\| = \|y\|$. Thus, the density remains $F(y)$. Since both random vectors have the same density (which is the product density), the components of AX are jointly independent as well. □

Corollary 1. *Let $X_1, X_2 : \Omega \to \mathbb{R}$ be i.i.d. RVs with $\mathcal{N}(0, \sigma^2)$. Then $X_1 + X_2$ and $X_1 - X_2$ are i.i.d. with $\mathcal{N}(0, 2\sigma^2)$.*

Proof Sketch. Consider the scaled vector $(\frac{X_1}{\sigma}, \frac{X_2}{\sigma})$ which is a standard Gaussian vector. Apply the rotation matrix:

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{X_1}{\sigma} \\ \frac{X_2}{\sigma} \end{bmatrix} = \begin{bmatrix} \frac{X_1+X_2}{\sigma\sqrt{2}} \\ \frac{X_1-X_2}{\sigma\sqrt{2}} \end{bmatrix}$$

The result is a Standard Gaussian vector. Thus, $\frac{X_1+X_2}{\sigma\sqrt{2}}$ and $\frac{X_1-X_2}{\sigma\sqrt{2}}$ are independent standard normals, implying X_1+X_2 and X_1-X_2 are independent with variance $(\sigma\sqrt{2})^2 = 2\sigma^2$. □

Proposition 1. *If X, Y are d-dimensional Gaussian random vectors such that $E(X) = E(Y)$ and $cov(X) = cov(Y)$, then $X \overset{d}{=} Y$ (they have the same distribution).*

Proof Sketch. Let $X^T = AX_1^T + b^T$ and $Y^T = BX_2^T + c^T$. We have $E(X) = b$ and $E(Y) = c$, so $b = c$. Assume $E(X) = 0$ for simplicity. Thus $X^T = AX_1^T$ and $Y^T = BX_2^T$. $A \in \text{Mat}(d \times m)$ and $B \in \text{Mat}(d \times k)$. WLOG assume $m \leq k$. By extending A and X_1^T by adding zeros, we may assume $A, B \in \text{Mat}(d \times k)$. Let $A_1, \dots, A_d$ be the rows of A and $B_1, \dots, B_d$ be the rows of B. Let $\mathcal{A} = \text{span}\{A_i\} \subset \mathbb{R}^k$ and $\mathcal{B} = \text{span}\{B_i\} \subset \mathbb{R}^k$. Define the map $L : \mathcal{A} \to \mathcal{B}$ by $L(A_i) = B_i$. We show L is injective. If $\sum v_i A_i = 0$, then:

$$\|\sum v_i A_i\|^2 = vAA^Tv^T = vBB^Tv^T = \|\sum v_i B_i\|^2 = 0$$

Thus L is an isomorphism and preserves the inner product (since $(AA^T)_{ij} = (BB^T)_{ij}$). L can be extended to an orthogonal transformation. Therefore Y^T can be written in terms of a transformed standard Gaussian vector that has the same distribution as X. □

Corollary 2. *Let Y be a Gaussian random vector ($Y^T = AX^T + b^T$). Then the components of Y are jointly independent if and only if its covariance matrix is diagonal (pairwise independent).*

Proof Sketch. If independent, the covariance matrix is clearly diagonal. Conversely, if $\mathrm{cov}(Y)$ is diagonal (say entries σ_i^2), consider a random vector Z with independent components $\mathcal{N}(0, \sigma_i^2)$. Z is Gaussian and has the same Mean and Covariance as Y. By the Proposition, $Y \stackrel{d}{=} Z$. Since Z has independent entries, so does Y. □

Brownian Motion

Theorem 1 (Wiener, 1923)**.** *The Standard Brownian Motion exists. That is, there exists a random process $B : [0, \infty) \times \Omega \to \mathbb{R}$ such that:*

1. *$B(0) = 0$ almost surely.*
2. *Independent increments: For $0 \le t_1 < t_2 < \cdots < t_k$, the increments $B(t_{i+1}) - B(t_i)$ are independent.*
3. *Stationary Gaussian increments: $B(t) - B(s) \sim \mathcal{N}(0, t - s)$ for $t > s$.*
4. *$t \mapsto B(t)$ is continuous almost surely.*

Construction of Brownian Motion on $[0, 1]$

Proof Sketch. We construct BM as a random element in $C[0, 1]$, as the uniform limit of a sequence of continuous functions defined on Dyadic Numbers. Let $\mathcal{D}_n = \left\{\frac{k}{2^n} : 0 \le k \le 2^n\right\}$. Note that $\mathcal{D}_n \subset \mathcal{D}_{n+1}$. Let $\mathcal{D} = \bigcup_{n \ge 0} \mathcal{D}_n$. We will define $B(d)$ for all $d \in \mathcal{D}$. Let $\{Z_d\}_{d \in \mathcal{D}}$ be independent $\mathcal{N}(0, 1)$ random variables.

Step $n = 0$: $\mathcal{D}_0 = \{0, 1\}$. Define $B(0) = 0$ and $B(1) = Z_1 \sim \mathcal{N}(0, 1)$. The increment is $B(1) - B(0) = Z_1$.

Step $n = 1$: $\mathcal{D}_1 = \{0, 1, \frac{1}{2}\}$. Define:

$$B\left(\frac{1}{2}\right) = \frac{B(1) + B(0)}{2} + \frac{Z_{1/2}}{2} = \frac{Z_1}{2} + \frac{Z_{1/2}}{2}$$

Increments:

- $B(1) - B(1/2) = Z_1 - (\frac{Z_1}{2} + \frac{Z_{1/2}}{2}) = \frac{Z_1}{2} - \frac{Z_{1/2}}{2}$. By Lemma, this is $\sim \mathcal{N}(0, \frac{1}{4} + \frac{1}{4}) = \mathcal{N}(0, \frac{1}{2})$.
- $B(1/2) - B(0) = \frac{Z_1}{2} + \frac{Z_{1/2}}{2}$. By Lemma, this is $\sim \mathcal{N}(0, \frac{1}{2})$.

Also, these increments are independent (sum and difference of independent Gaussians).

Step $n = 2$: $\mathcal{D}_2 = \{0, 1, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}\}$.

$$B\left(\frac{1}{4}\right) = \frac{B(1/2) + B(0)}{2} + \frac{Z_{1/4}}{2^2}$$

$$B\left(\frac{3}{4}\right) = \frac{B(1)+B(1/2)}{2} + \frac{Z_{3/4}}{2^2}$$

Let us check the increments. First, for the interval $[0, \frac{1}{4}]$:

$$B\left(\frac{1}{2^2}\right) - B(0) = \frac{Z_1}{4} + \frac{Z_{1/2}}{4} + \frac{Z_{1/2^2}}{2^{3/2}} \sim \mathcal{N}\left(0, \frac{1}{4}\right)$$

Next, for the interval $[\frac{1}{4}, \frac{1}{2}]$:

$$B\left(\frac{1}{2}\right) - B\left(\frac{1}{2^2}\right) = \frac{Z_1}{2} + \frac{Z_{1/2}}{2} - \left(\frac{Z_1}{4} + \frac{Z_{1/2}}{4} + \frac{Z_{1/2^2}}{2^{3/2}}\right) = \frac{Z_1}{4} + \frac{Z_{1/2}}{4} - \frac{Z_{1/2^2}}{2^{3/2}}$$

By Lemma, those two increments are i.i.d with $\mathcal{N}\left(0, \frac{1}{4}\right)$.

Now consider the wider interval increment $B(1) - B\left(\frac{1}{2^2}\right)$:

$$B(1) - B\left(\frac{1}{2^2}\right) = Z_1 - \left(\frac{Z_1}{4} + \frac{Z_{1/2}}{4} + \frac{Z_{1/2^2}}{2^{3/2}}\right) = \frac{3Z_1}{4} - \frac{Z_{1/2}}{4} - \frac{Z_{1/2^2}}{2^{3/2}}$$

Each of the terms have a different variance but the total variance is $\frac{3}{4}$.

Finally, consider $B(\frac{3}{2^2}) - B(0)$:

$$B\left(\frac{3}{2^2}\right) - B(0) = \frac{3Z_1}{4} + \frac{Z_{1/2}}{4} + \frac{Z_{3/2^2}}{2^{3/2}}$$

This is the same as above with variance $\frac{3}{4}$. The independence of those increments follows from the fact that they are two linear combinations of 3 independent terms.

General Step: Inductively define $B(d)$ for $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$.

$$B(d) = \frac{B(d+\frac{1}{2^n}) + B(d-\frac{1}{2^n})}{2} + \frac{Z_d}{2^{(n+1)/2}}$$

This ensures increments are pairwise (and jointly) independent with the correct distribution.

Functional Representation

Let us define functions $F_n(t)$.

$$F_0(t) = \begin{cases} 0 & t=0 \\ Z_1 & t=1 \\ \text{Linear} & \text{in between} \end{cases} = tZ_1$$

$$F_1(t) = \begin{cases} 0 & t \in \mathcal{D}_0 \\ \frac{Z_{1/2}}{2} & t = \frac{1}{2} \\ \text{Linear} & \text{in between} \end{cases}$$

$$F_n(t) = \begin{cases} 0 & t \in \mathcal{D}_{n-1} \\ \frac{Z_d}{2^{(n+1)/2}} & t \in \mathcal{D}_n \setminus \mathcal{D}_{n-1} \\ \text{Linear} & \text{in between} \end{cases}$$

Define the partial sum:

$$S_n(d) = \sum_{i=0}^{n} F_i(d)$$

Since for $d \in \mathcal{D}_n$, $F_m(d) = 0$ for all $m > n$, the sum stabilizes. We estimate the limit $S(t) = \sum_{n=0}^{\infty} F_n(t)$ for $t \in [0, 1]$. If this sum converges uniformly, the Brownian Motion is well-defined and continuous.

□