Chapter - 5: Random Vectors

Aveek Dutta
Assistant Professor
Department of Electrical and Computer Engineering
University at Albany
Fall 2019

PDF/CDF

- Groups of RVs are studied in form of vectors called ***random vectors***
  - **Bold Uppercase** letters denote random vectors and matrices
  - **Bold Lowercase** letters denote deterministic vectors, e.g., values a random vector assumes
- Event ζ is mapped to the real line by multiple RVs, $X_1, X_2, \ldots, X_N$, which forms and N-dimensional vector $\boldsymbol{X}(\zeta) \triangleq [X_1(\zeta), X_2(\zeta), \ldots, X_N(\zeta)] \in R^n$
  - $\boldsymbol{X}$ can be real or imaginary.
- Therefore, the CDF of a random vector is

$$F_{\mathbf{X}}(\mathbf{x}) \triangleq P\left[X_1 \leq x_1, \ldots, X_n \leq x_n\right]$$

- By defining $\{\mathbf{X} \leq \mathbf{x}\} \triangleq \{X_1 \leq x_1, \ldots, X_n \leq x_n\}$, we can write $F_{\mathbf{X}}(\mathbf{x}) \triangleq P[\mathbf{X} \leq \mathbf{x}]$
  - Also, certain and impossible event and the pdf can be written as

$$F_{\mathbf{X}}(\infty) = 1 \qquad F_{\mathbf{X}}(-\infty) = 0$$

$$f_{\mathbf{X}}(\mathbf{x}) \triangleq \frac{\partial^n F_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \ldots \partial x_n}$$

- Also observe that

$$f_{\mathbf{X}}(\mathbf{x}) \Delta x_1 \ldots \Delta x_n \simeq P\left[x_1 < X_1 \leq x_1 + \Delta x_1, \ldots, x_n < X_n \leq x_n + \Delta x_n\right] \longrightarrow$$

$$F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\mathbf{X}}(\mathbf{x}')\, dx_1' \ldots dx_n'$$

$$F_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{\mathbf{x}} f_{\mathbf{X}}(\mathbf{x}')\, d\mathbf{x}'$$

More generally, with $B \subset R^N$

$$P[B] = \int_{\mathbf{x} \in B} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

Conditional and marginals

- Conditional CDF of **X** given *B*

$$F_{\mathbf{X}|B}(\mathbf{x}|B) \triangleq P[\mathbf{X} \le \mathbf{x}|B] = \frac{P[\mathbf{X} \le \mathbf{x}, B]}{P[B]} \quad (P[B] \neq 0) \iff F_{\mathbf{X}}(\mathbf{x}) = \sum_{i=1}^{n} F_{\mathbf{X}|B_i}(\mathbf{x}|B_i)P[B_i]$$

- Conditional densities are given by

$$f_{\mathrm{X}|B}(\mathrm{x}|B) \triangleq \frac{\partial^n F_{\mathrm{X}|B}(\mathrm{x}|B)}{\partial x_1 \ldots \partial x_n} \longrightarrow f_{\mathbf{X}}(\mathbf{x}) = \sum_{i=1}^{n} f_{\mathbf{X}|B}(\mathbf{x}|B_i)P[B_i]$$

- Joint distribution and densities -

$$F_{\mathrm{XY}}(\mathrm{x},\mathrm{y}) = P[\mathrm{X} \le \mathrm{x}, \mathrm{Y} \le \mathrm{y}] \longrightarrow f_{\mathbf{XY}}(\mathbf{x},\mathbf{y}) = \frac{\partial^{(n+m)} F_{\mathbf{XY}}(\mathbf{x},\mathbf{y})}{\partial x_1 \ldots \partial x_n \partial y_1 \ldots \partial y_m}$$

- Marginal pdf by integrating joint pdf

$$f_{\mathbf{X}}(\mathrm{x}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{XY}}(\mathrm{x},\mathrm{y})dy_1 \ldots dy_m$$

- Marginals for $\mathbf{X}' \triangleq (X_1, \ldots, X_{n-1})^T$

$$f_{\mathbf{X}'}(\mathbf{x}') \triangleq \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})dx_n \quad \text{where } \mathbf{x}' \triangleq (x_1, \ldots, x_{n-1})^T$$

Integration over x_n gives marginal density of x’

Multiple Transformation of RV

- If $\boldsymbol{X} \triangleq [X_1, X_2, \ldots, X_N]$ is a random vector and define n *functionally independent* real functions as another rv $\boldsymbol{Y} \triangleq [Y_1, Y_2, \ldots, Y_N]$,

$$\begin{aligned} y_1 &= g_1(x_1, x_2, \cdots, x_n) \\ y_2 &= g_2(x_1, x_2, \cdots, x_n) \\ &\vdots \\ y_n &= g_n(x_1, x_2, \cdots, x_n) \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} x_1 &= \phi_1(y_1, y_2, \cdots, y_n) \\ x_2 &= \phi_2(y_1, y_2, \cdots, y_n) \\ &\vdots \\ x_n &= \phi_n(y_1, y_2, \cdots, y_n) \end{aligned}$$

- Following the discussion in case of transformation of two rv in chap - 3, except the infinitesimal hypervolume is defined in n-dimensional space,

$$P[A] = \int_{\mathscr{P}_y} f_Y(\mathbf{y})dy = f_{\mathbf{Y}}(\mathbf{y})\mathrm{V}_y = \int_{\mathscr{P}_x} f_{\mathbf{X}}(\mathbf{x})dx = f_{\mathbf{X}}(\mathbf{x})\mathrm{V}_x$$

Elementary event that defines the hypervolume V_y

$$A \triangleq \{\zeta : y_i \le Y_i \le y_i + dy_i, i = 1, \ldots, n\}$$

Ratio is defined by the determinant of the Jacobian

Recall: Two variable case

$$\begin{aligned} P[v < V \le v + dv, w < W \le w + dw] &= \iint_{\mathscr{R}} f_{VW}(\xi, \eta) d\xi d\eta \\ &= f_{VW}(v, w) A(\mathscr{R}) \\ &= \iint_{\mathscr{S}} f_{XY}(\xi, \eta) d\xi d\eta \\ &= f_{XY}(x, y) A(\mathscr{S}) \end{aligned}$$

- Jacobian determinant is given by

$$\tilde{J} = \begin{vmatrix} \frac{\partial \phi_1}{\partial y_1} & \cdots & \frac{\partial \phi_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial \phi_n}{\partial y_1} & \cdots & \frac{\partial \phi_n}{\partial y_n} \end{vmatrix} = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{vmatrix}^{-1} = J^{-1}$$

$$\boxed{f_{\mathbf{X}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{x})|\tilde{J}| = f_{\mathbf{X}}(\mathbf{x})/|J|}$$

If $\boldsymbol{r}$ roots exists for **Y**=g(**X**), then multiple Jacobians define the ratio of disjoint hypervolumes,

$$f_{\mathbf{Y}}(\mathbf{y}) = \sum_{i=1}^{r} f_{\mathbf{X}}\left(\mathbf{x}^{(i)}\right)\left|\tilde{J}_i\right| = \sum_{i=1}^{r} f_{\mathbf{X}}\left(\mathbf{x}^{(i)}\right)/|J_i|$$

where, $\left|\tilde{J}_i\right| \triangleq \mathrm{V}_x^{(i)}/\mathrm{V}_y$

Example 5.2-1

- We are given three functions

$$g_1(\mathbf{x}) = x_1^2 - x_2^2 \quad \text{OR} \quad y_1 = x_1^2 - x_2^2$$
$$g_2(\mathbf{x}) = x_1^2 + x_2^2 \longrightarrow y_2 = x_1^2 + x_2^2$$
$$g_3(\mathbf{x}) = x_3 \qquad y_3 = x_3$$

Given

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-3/2} \exp\left[-\frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right)\right]$$

It has four solutions, with four disjoint hypervolumes

$$x_1^{(1)} = ((y_1 + y_2)/2)^{1/2} \qquad x_1^{(2)} = ((y_1 + y_2)/2)^{1/2}$$
$$x_2^{(1)} = ((y_2 - y_1)/2)^{1/2} \qquad x_2^{(2)} = -((y_2 - y_1)/2)^{1/2}$$
$$x_3^{(1)} = y_3 \qquad x_3^{(2)} = y_3$$

$$x_1^{(3)} = -((y_1 + y_2)/2)^{1/2} \qquad x_1^{(4)} = -((y_1 + y_2)/2)^{1/2}$$
$$x_2^{(3)} = ((y_2 - y_1)/2)^{1/2} \qquad x_2^{(4)} = -((y_2 - y_1)/2)^{1/2}$$
$$x_3^{(3)} = y_3 \qquad x_3^{(4)} = y_3$$

- The Jacobian is given by $J = \begin{vmatrix} 2x_1 & -2x_1 & 0 \\ 2x_1 & 2x_2 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 8x_1x_2$
- Note, $|J_1| = |J_2| = |J_3| = |J_4| = 4\left(y_2^2 - y_1^2\right)^{1/2}$
- Also note, for roots to be real,

$$y_2 \geq 0, y_1 + y_2 \geq 0, \text{ and } y_2 - y_1 \geq 0. \text{ Hence } y_2 \geq |y_1|$$

- Therefore,

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{4\left(y_2^2 - y_1^2\right)^{1/2}} \sum_{i=1}^{4} f_{\mathbf{X}}(\mathbf{x}_i) = \frac{(2\pi)^{-3/2}}{\left(y_2^2 - y_1^2\right)^{1/2}} \exp\left[-\frac{1}{2}\left(y_2 + y_3^2\right)\right] \times u(y_2)u(y_2 - |y_1|)$$

Expectation Vectors

- **Definition:** The expected value of the column vector **X** = $[X_1, X_2, \ldots, X_N]^T$ is a vector $\boldsymbol{\mu}$ whose elements $\mu_1, \mu_2, \ldots, \mu_N$, are given by

$$\mu_i \triangleq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i f_{\mathbf{X}}(x_1, \ldots, x_n) dx_1 \ldots dx_n$$

- Alternately, using marginal density of $\mathbf{X_i}$

Integrate over all indices except for *i*

$$\mu_i = \int_{-\infty}^{\infty} x_i f_{X_i}(x_i) dx_i \quad i = 1, \ldots, n \qquad f_{X_i}(x_i) \triangleq \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n$$

- **Definition:** Covariance matrix **K** is the vector outer product

$$\mathbf{K} \triangleq E\big[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T\big] \longrightarrow \begin{aligned} K_{ij} &\triangleq E[(X_i - \mu_i)(X_j - \mu_j)] \\ &= E[(X_j - \mu_j)(X_i - \mu_i)] \\ &= K_{ji} \quad i, j = 1, \ldots, n \end{aligned}$$

- **Define** $\sigma_i^2 \triangleq K_{ii}$

$$\mathbf{K} = \begin{bmatrix} \sigma_1^2 & & \cdots & & K_{1n} \\ & \ddots & & & \\ \vdots & & \sigma_i^2 & & \vdots \\ & & & \ddots & \\ K_{n1} & & \cdots & & \sigma_n^2 \end{bmatrix}$$

- Def: Correlation matrix $\mathbf{R} \triangleq E[\mathbf{X}\mathbf{X}^T]$
  - Expanding the covariance matrix $\mathbf{K} = \mathbf{R} - \boldsymbol{\mu}\boldsymbol{\mu}^T$
- Def: Vectors **X** and **Y** Uncorrelated if $E\{\mathbf{X}\mathbf{Y}^T\} = \mu_{\mathbf{X}} \boldsymbol{\mu}_{\mathbf{Y}}^T$
- Def: Orthogonal if $E\{\mathbf{X}\mathbf{Y}^T\} = 0$

Example 5.4-1

- Given $f_{X_1X_2}(x_1, x_2) = x_1 + x_2$ for $0 < x_1 \leq 1, 0 < x_2 \leq 1$, Compute $\mathbf{K} \triangleq E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T]$

Solution: We know, $K_{12} = K_{21} = R_{21} - \mu_2\mu_1$

By Definition:

$$\mu_1 = \mu_2 = \iint_S x(x+y)dxdy = 0.583$$

$$R_{12} = R_{21} \triangleq \iint_S xy(x+y)dxdy = 0.333$$

where $S = \{(x_1, x_2) : 0 < x_1 \leq 1, 0 < x_2 \leq 1\}$

$$\boxed{\text{Hence } K_{12} = K_{21} = 0.333 - (0.583)^2 = -0.007}$$

The diagonals of the covariance matrix is given by $\sigma^2 = E\left[X^2\right] - \mu^2$

$$\sigma_1^2 = \sigma_2^2 = \int_0^1 x^2(x + 1/2)dx - (0.583)^2 = 0.077$$

$$\boxed{\mathbf{K} = \begin{bmatrix} 0.077 & -0.007 \\ -0.007 & 0.077 \end{bmatrix} = 0.077 \begin{bmatrix} 1 & -0.09 \\ -0.09 & 1 \end{bmatrix}}$$

Properties of Covariance Matrix

- Covariance matrix is ***at least positive semi-definite***
  - For any column vector **z** and real symmetric matrix **M** is p.s.d. if,

$$q(\mathbf{z}) \triangleq \mathbf{z}^T \mathbf{M} \mathbf{z} \geq 0$$

A scalar

Positive definite if >0

General example:
And for a r.s matrix M $= \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$
Any vector z = [a b c]^T

$$\begin{aligned} z^{\mathsf T} M z = \left(z^{\mathsf T} M\right) z &= \left[\,(2a-b) \quad (-a+2b-c) \quad (-b+2c)\,\right] \begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= (2a-b)a + (-a+2b-c)b + (-b+2c)c \\ &= 2a^2 - ba - ab + 2b^2 - cb - bc + 2c^2 \\ &= 2a^2 - 2ab + 2b^2 - 2bc + 2c^2 \\ &= a^2 + a^2 - 2ab + b^2 + b^2 - 2bc + c^2 + c^2 \\ &= a^2 + (a-b)^2 + (b-c)^2 + c^2 \end{aligned}$$

This is a sum of squares of real numbers, whis is always ≥ 0

- For any covariance matrix $\mathbf{K}_{\mathbf{XX}}$ and any vector **z**, define $\mathbf{Y} = \mathbf{z}^T\mathbf{X}$ (a scalar),

$$\begin{aligned} 0 \leq \mathrm{Var}(Y) = \mathrm{Cov}(Y) &= E[(Y-\mu)(Y-\mu)^T] \\ &= E\left[\mathbf{z}^T(\mathbf{X}-\boldsymbol{\mu}_X)(\mathbf{X}-\boldsymbol{\mu}_X)^T\mathbf{z}\right] \\ &= \mathbf{z}^T E\left[(\mathbf{X}-\boldsymbol{\mu}_X)(\mathbf{X}-\boldsymbol{\mu}_X)^T\right]\mathbf{z} \\ &= \mathbf{z}^T \mathbf{K}_{\mathbf{XX}} \mathbf{z}, \quad \mathbf{K}_{\mathbf{XX}} \triangleq E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T\right] \end{aligned}$$

- Eigenvalues and eigenvectors of M can also be calculated for **K.**
  - Eigenvalues (λ) are solutions to $\det(\mathbf{M} - \lambda\mathbf{I}) = 0$
  - Corresponding eigenvectors are obtained by solving **(M- λI)ϕ = 0**
  - **See example 5.4-1**

$$\mathbf{M} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix} \longrightarrow \det\begin{bmatrix} 4-\lambda & 2 \\ 2 & 4-\lambda \end{bmatrix} = (4-\lambda)^2 - 4 = 0 \longrightarrow (\mathbf{M} - 6\mathbf{I})\phi = 0$$

$$\lambda_1 = 6, \quad \lambda_2 = 2$$

Definitions

- A & B are similar matrices if there exists a n x n matrix T with det(T) ≠ 0, s.t. $\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \mathbf{B}$
- **Theorem:** An *n x n* matrix **M** is <u>similar</u> to a diagonal matrix iff **M** has linearly independent eigenvectors.
- **Theorem:** If **M** is a r.s matrix with eigenvalues $\lambda_1, \lambda_2, \ldots$, then M has n mutually orthogonal **unit** eigenvectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots \boldsymbol{\phi}_N$.

- From the two Theorems, if **M is r.s and has orthogonal (and therefore linearly independent) eigenvectors** then it is **similar** to a diagonal matrix **Λ** under some transformation **T.**
  - So under the transformation $\mathbf{U}^{-1}\mathbf{M}\mathbf{U} = \boldsymbol{\Lambda}$, with $\mathbf{U} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots \boldsymbol{\phi}_N]$ and $\Lambda \triangleq \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$
  - Since $\mathbf{U}^T\mathbf{U} = 1$ and $\mathbf{U}^T = \mathbf{U}^{-1}$ **(unitary matrices),** we can write $\mathbf{U}^T\mathbf{M}\mathbf{U} = \boldsymbol{\Lambda}$
- Distance preserving property under the transformation **y=Ux**

$$\|\mathbf{y}\|^2 = \mathbf{y}^T\mathbf{y} = \mathbf{x}^T\mathbf{U}^T\mathbf{U}\mathbf{x} = \|\mathbf{x}\|^2$$

- Key Takeaway - **K** can be diagonalized if eigenvectors (**U**) are known.
  - Why is it useful to diagonalize?

Properties of K_{xx}

- **Theorem:** Iff all eigenvalues are positive then a r.s matrix M is positive definite
- **Proof:** Let λ_i>0, then for any column vector and transformation **x=Uy** we can write,

$$\begin{aligned}\mathbf{x}^T\mathbf{M}\mathbf{x} &= (\mathbf{U}\mathbf{y})^T\mathbf{M}(\mathbf{U}\mathbf{y})\\ &=\mathbf{y}^T\mathbf{U}^T\mathbf{M}\mathbf{U}\mathbf{y}\\ &=\mathbf{y}^T\mathbf{\Lambda}\mathbf{y}\\ &=\sum_{i=1}^{n}\lambda_i y_i^2 > 0\end{aligned}$$

Λ is the diagonal matrix

Unless **y**=0, in that case **x**= 0

- Conversely, replace **x** by $\boldsymbol{\phi}_i$, if M is p.d. then λ_i>0,

$$0 < \mathbf{x}^T\mathbf{M}\mathbf{x}$$

$$0 < \phi_i^T\mathbf{M}\phi_i = \lambda_i \quad i = 1, \ldots, n$$

If M is p.d then eigenvalues are also positive.

Whitening Transform

- Given a zero mean n x 1 random vector **X,** with **p.d.** Covariance matrix $\mathbf{K_{XX}}$. Find a transformation **Y=CX** such that $\mathbf{K_{YY}} = \mathbf{I}$.
  - **C** is called the whitening matrix and the transformation is called whitening.
- Solution: The characteristic equation $(\mathbf{K_{XX}} - \lambda_i \mathbf{I})\phi_i = 0$ or $\mathbf{K_{XX}}\phi_i = \lambda_i \phi_i$ can be written as $\mathbf{K_{XX} U = U \Lambda}$ (U is the eigenvector matrix, $\mathbf{\Lambda}$ is the diagonal eigenvalue matrix.
  - Since K_{XX} is p.d., $\lambda_i > 0$ and therefore, $\mathbf{\Lambda}^{1/2} = \text{diag}[1\sqrt{\lambda_i}, 1\sqrt{\lambda_2} \ldots 1\sqrt{\lambda_n}]$ is also well defined.
  - Consider the transform $\mathbf{Y = CX = \Lambda^{-1/2} U^T X}$, we get

$$\mathrm{K_{YY}} = E[\mathrm{YY}^T] = E[\mathrm{CXX}^T\mathrm{C}^T] = \Lambda^{-1/2}\mathrm{U}^T E[\mathrm{XX}^T]\mathrm{U}\Lambda^{-1/2} = \Lambda^{-1/2}\mathrm{U}^T\mathrm{K_{XX}}\mathrm{U}\Lambda^{-1/2}$$

$$= \Lambda^{-1/2}\mathrm{U}^T(\mathrm{K_{XX}}\mathrm{U})\Lambda^{-1/2} = \Lambda^{-1/2}\mathrm{U}^T(\mathrm{U}\Lambda)\Lambda^{-1/2} = \Lambda^{-1/2}(\mathrm{U}^T\mathrm{U})\Lambda\Lambda^{-1/2} = \Lambda^{-1/2}\Lambda\Lambda^{-1/2} = \mathrm{I}, \text{ since } \mathrm{U}^T\mathrm{U} = \mathrm{I}$$

$$\mathbf{K_{XX}} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \quad \mathbf{U} = \begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/2 & 1/2 \\ 1/\sqrt{2} & 1/2 & -1/2 \end{bmatrix} = \mathbf{U}^T \quad \mathbf{\Lambda} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2+\sqrt{2} & 0 \\ 0 & 0 & 2-\sqrt{2} \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} 1/\sqrt{2} & 0 & 0 \\ 0 & (2+\sqrt{2})^{-1/2} & 0 \\ 0 & 0 & (2-\sqrt{2})^{-1/2} \end{bmatrix} \begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/-2 & 1/2 \\ 1/\sqrt{2} & 1/2 & -1/2 \end{bmatrix} \mathbf{X}$$

Example 5.5-2

- Given $\mathbf{K_X} = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$. Find transform **U** to form **Y** with diagonal $\mathbf{K_{YY}}$
- Start with eigenvalues by solving det($\mathbf{K_{XX}}$- λ**I)** = 0
  - yields $\lambda_1 = 2$, $\lambda_2 = 2+\sqrt{2}$, $\lambda_3 = 2-\sqrt{2}$
- Compute the the three (normalized) orthogonal eigenvectors $(\mathbf{K_{XX}} - \lambda_i \mathbf{I})\boldsymbol{\phi}_i = 0$

$$\phi_1 = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T \qquad \phi_2 = \left(\frac{1}{\sqrt{2}}, -\frac{1}{2}, \frac{1}{2}\right)^T \qquad \phi_3 = \left(\frac{1}{\sqrt{2}}, \frac{1}{2}, -\frac{1}{2}\right)^T$$

- Create eigenvector matrix $\mathbf{U}^T = [\boldsymbol{\phi}_1\, \boldsymbol{\phi}_2\, \boldsymbol{\phi}_3]^T$

$$\mathbf{A} = \mathbf{U}^T = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

- Now, the transformation **Y = AX** yields

$$\begin{aligned} Y_1 &= \tfrac{1}{\sqrt{2}}(X_2 + X_3) \\ Y_2 &= \tfrac{1}{\sqrt{2}}X_1 - \tfrac{1}{2}X_2 + \tfrac{1}{2}X_3 \\ Y_3 &= \tfrac{1}{\sqrt{2}}X_1 + \tfrac{1}{2}X_2 - \tfrac{1}{2}X_3 \end{aligned}$$

And $\mathbf{K_{YY}} = \mathbf{U^{-T} M U}$

$$\mathbf{K_Y} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2+\sqrt{2} & 0 \\ 0 & 0 & 2-\sqrt{2} \end{bmatrix}$$

Multidimensional Gaussian

Scalar RV X

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

Random Vector
$\mathbf{X}=[X_1,X_2,....X_n]^T$ with
independent components

$$f_{\mathbf{X}}(x_1,\ldots,x_n) = \prod_{i=1}^{n} f_{X_i}(x_i) = \frac{1}{(2\pi)^{n/2}\sigma_1\ldots\sigma_n}\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2\right]$$

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}[\det(\mathbf{K})]^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{K}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

$$\mathbf{K} \triangleq \begin{bmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2\end{bmatrix} \qquad \mu = (\mu_1,\ldots,\mu_n)^T, \text{ and } \det(\mathbf{K}) = \prod_{i=1}^{n}\sigma_i^2.$$

- Is $f_X(\mathbf{x})$ a pdf for **any** arbitrary p.d matrix $\mathbf{K}_{XX}$. We have to prove $\int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = 1$
  - Define $\mathbf{z} = \mathbf{x} - \boldsymbol{\mu}$, then the pdf can be written as

$$\phi(\mathbf{z}) \triangleq \exp\left(-\tfrac{1}{2}\mathbf{z}^T\mathbf{K}^{-1}\mathbf{z}\right)$$
$$\alpha \triangleq \int_{-\infty}^{\infty}\phi(\mathbf{z})d\mathbf{z}$$
$$\int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} = \frac{\alpha}{(2\pi)^{n/2}[\det(\mathbf{K})]^{1/2}}$$

→ Under $\mathbf{z}=\mathbf{C}\mathbf{y}$ transform and noting $\mathbf{K}_{XX} = \mathbf{C}\mathbf{C}^T$ and $\mathbf{C}^T\mathbf{K}_{XX}^{-1}\mathbf{C} = \mathbf{I}$ →

$$\mathbf{z}^T\mathbf{K_{XX}}^{-1}\mathbf{z} = \mathbf{y}^T\mathbf{C}^T\mathbf{K_{XX}}^{-1}\mathbf{C}\mathbf{y} = \|\ \mathbf{y}\ \|^2 = \sum_{i=1}^{n} y_i^2$$

so that $\phi(\mathbf{z})$ is given by

$$\phi(\mathbf{z}) = \prod_{i=1}^{n}\exp\left[-\tfrac{1}{2}y_i^2\right]$$

But this transformer matrix C matrix will have to be a special matrix (see next page)

Proof for $K_{XX} = CC^T$ and $C^T K_{XX}^{-1} C = I$

From whitening discussion we have

$$\mathbf{K_{YY}} = \mathbf{\Lambda}^{-1/2}\mathbf{U}^T\mathbf{K_{XX}}\mathbf{U}\mathbf{\Lambda}^{-1/2} = \mathbf{I}$$
$$\text{let } \mathbf{\Lambda}^{-1/2} = \mathbf{Z} = [\mathbf{\Lambda}^{-1/2}]^T$$
$$= (\mathbf{UZ})^T\mathbf{K_{XX}}(\mathbf{UZ}) = \mathbf{I}$$

NOTE:
(i) $U^T = U^{-1}$
(ii) $Z^T = Z \rightarrow [Z^{-1}]^T = Z^{-1}$
(iii) $[A^T]^{-1} = [A^{-1}]^T$

Pre-multiply by $[(\mathbf{UZ})^T]^{-1}$ and post-multiply by $(\mathbf{UZ})^{-1}$ to isolate $\mathbf{K_{XX}}$, we get

$$\mathbf{K_{XX}} = [(\mathbf{UZ})^T]^{-1} . [\mathbf{UZ}]^{-1} = (\mathbf{Z}^T \mathbf{U}^T)^{-1} . \mathbf{Z}^{-1}\mathbf{U}^{-1}$$

$$= (\mathbf{U}^T)^{-1} . (\mathbf{Z}^T)^{-1} . \mathbf{Z}^{-1}\mathbf{U}^{-1} = (\mathbf{U}^{-1})^{-1} . (\mathbf{Z}^{-1})^T . \mathbf{Z}^{-1}\mathbf{U}^{-1} = \mathbf{UZ}^{-1} . \mathbf{Z}^{-1}\mathbf{U}^{-1}$$

$$= \mathbf{UZ}^{-1} . [\mathbf{Z}^{-1}]^T \mathbf{U}^T = \mathbf{UZ}^{-1} . [\mathbf{UZ}^{-1}]^T = \mathbf{CC}^T \text{ (where, } \mathbf{C} = \mathbf{UZ}^{-1})$$

- Also, $\mathbf{C}^T \mathbf{K_{XX}}^{-1} \mathbf{C} = \mathbf{C}^T [\mathbf{CC}^T]^{-1}\mathbf{C} = \mathbf{C}^T [\mathbf{C}^T]^{-1}\mathbf{C}^{-1}\mathbf{C} = (\mathbf{C}^{-1}\mathbf{C})^T (\mathbf{C}^{-1}\mathbf{C}) = \mathbf{I}$
- For any p.d. matrix **P**, there exists **C** such that $\mathbf{P} = \mathbf{CC}^T$ and $\mathbf{C}^T\mathbf{K_{XX}}^{-1}\mathbf{C} = \mathbf{I}$

...contd

Volume elements are related as below for a linear transformation **z** = **Cy**

$$d\mathbf{z} = |\det(\mathbf{C})| d\mathbf{y} \qquad \text{where } d\mathbf{z} \triangleq dz_1 \dots dz_n \text{ and } d\mathbf{y} = dy_1 \dots dy_n$$

Theorem (Determinants and volumes). *Let $v_1, v_2, \dots, v_n$ be vectors in $\mathbf{R}^n$, let P be the parallelepiped determined by these vectors, and let A be the matrix with rows $v_1, v_2, \dots, v_n$. Then the absolute value of the determinant of A is the volume of P:*

$$|\det(A)| = \mathrm{vol}(P).$$

Reference: https://textbooks.math.gatech.edu/ila/determinants-volumes.html

Therefore, the integral reduces to

$$\alpha = \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} y_i^2\right) dy_1 \dots dy_n |\det(\mathbf{C})|$$

$$= \left[\int_{-\infty}^{\infty} e^{-y^2/2} dy\right]^n |\det(\mathbf{C})|$$

$$= [2\pi]^{n/2} |\det(\mathbf{C})|$$

Integral of the standard normal multiplied by √2π

$$\text{But since } \mathbf{K} = \mathbf{C}\mathbf{C}^T, \det(\mathbf{K}) = \det(\mathbf{C})\det\left(\mathbf{C}^T\right) = [\det(\mathbf{C})]^2$$

$$|\det(\mathbf{C})| = |\det(\mathbf{K})|^{1/2} = (\det(\mathbf{K}))^{1/2}$$

→ Since, det(AB) = det(A).det(B) and det(A) = det(A^T)

Therefore, we obtain,

$$\alpha = (2\pi)^{n/2}[\det(\mathbf{K})]^{1/2}$$

$$\Rightarrow \int_{-\infty}^{\infty} f\mathbf{x}(\mathbf{x}) d\mathbf{x} = \frac{\alpha}{(2\pi)^{n/2}[\det(\mathbf{K})]^{1/2}} = 1$$

Hence the multidimensional Gaussian pdf integrates to 1. This proves that it is a valid pdf.

Transformation of Gaussian pdf

Theorem: Let **X** be an n-dimensional Normal random vector with positive definite cov. Matrix $\mathbf{K}_{\mathbf{XX}}$ and mean vector $\boldsymbol{\mu}$. Let **A** be a nonsingular linear transformation in *n* dimensions. Then **Y** = **AX** is an *n*-dimensional Normal random vector with covariance matrix $\mathbf{K}_{\mathbf{YY}} = \mathbf{A}\mathbf{K}_{\mathbf{XX}}\mathbf{A}^{\mathbf{T}}$ and mean vector $\boldsymbol{\beta}=\mathbf{A}\boldsymbol{\mu}$.

Proof: Start with the Jacobian $f_{\mathbf{Y}}(\mathbf{y}) = \sum_{i=1}^{r} \frac{f_{\mathbf{X}}(\mathbf{x}_i)}{|J_i|}$ where, $\mathbf{Y} = \mathbf{g}(\mathbf{X}) \triangleq (g_1(\mathbf{X}), \dots, g_n(\mathbf{X}))^T$

The i^{th} Jacobian is

$$J_i = \det\left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right)\bigg|_{\mathbf{x}=\mathbf{x}_i} = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{vmatrix}_{\mathbf{x}=\mathbf{x}_i}$$

Since **A** is a non-singular linear transformation, the only solution of

$$\mathbf{A}\mathbf{x} - \mathbf{y} = 0 \quad \text{is} \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{y} \longrightarrow J_i = \det\left(\frac{\partial(\mathbf{A}\mathbf{x})}{\partial \mathbf{x}}\right) = \det(\mathbf{A})$$

which leads to

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}[\det(\mathbf{K})]^{1/2}|\det(\mathbf{A})|}\exp\left(-\frac{1}{2}(\mathbf{A}^{-1}\mathbf{y} - \boldsymbol{\mu})^T\mathbf{K}^{-1}(\mathbf{A}^{-1}\mathbf{y} - \boldsymbol{\mu})\right)$$

...contd

Now, $[\det(\mathbf{K})]^{1/2}|\det(\mathbf{A})| = \left[\det\left(\mathbf{AKA}^T\right)\right]^{1/2} \triangleq [\det(\mathbf{Q})]^{1/2}$

Since, det(A^T) - det(A) we get,
det(K).|det(A)|^2= det(A).det(K) .det(A)
= det(A).det(K) .det(A^T) = det(AKA^T)

Also, factoring out $\mathbf{A}^{-1}$ we get,

$$\begin{aligned}(\mathbf{A}^{-1}\mathbf{y}-\boldsymbol{\mu})^T\mathbf{K}^{-1}(\mathbf{A}^{-1}\mathbf{y}-\boldsymbol{\mu}) &= \left[\mathbf{A}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\mu})\right]^T\mathbf{K}^{-1}\mathbf{A}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\mu}) \\ &= (\mathbf{y}-\mathbf{A}\boldsymbol{\mu})^T[\mathbf{A}^{-1}]^T\mathbf{K}^{-1}\mathbf{A}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\mu}) \\ &= (\mathbf{y}-\mathbf{A}\boldsymbol{\mu})^T[\mathbf{A}^T]^{-1}\mathbf{K}^{-1}\mathbf{A}^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\mu}) \\ &= (\mathbf{y}-\mathbf{A}\boldsymbol{\mu})^T\left(\mathbf{AKA}^T\right)^{-1}(\mathbf{y}-\mathbf{A}\boldsymbol{\mu})\end{aligned}$$

Since, $(ABC)^{-1} = C^{-1}BA^{-1}$ and $[A^T]^{-1} = [A^{-1}]^T$
We get $[A^{-1}]^T K_{XX}^{-1} A^{-1} = [A^{-T}]^{-1} K_{XX}^{-1} A^{-1} = [AK_{XX} A^T]^{-1}$

Therefore, since, $\boldsymbol{\beta}=\mathbf{A}\mu$ and $\mathbf{Q} = \mathbf{E[(Y-\boldsymbol{\beta})(Y-\boldsymbol{\beta})^T]}$, we can write

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}[\det(\mathbf{Q})]^{1/2}}\exp\left[-\frac{1}{2}(\mathbf{y}-\boldsymbol{\beta})^T\mathbf{Q}^{-1}(\mathbf{y}-\boldsymbol{\beta})\right]$$

Theorem: The above holds for $\mathbf{A}_{\mathbf{mxn}}$ transformations as well

$$\begin{aligned}\mathrm{Q} &\triangleq \mathrm{A}_{mn}\mathrm{K}\mathrm{A}_{mn}^T \\ \beta &= \mathrm{A}_{mn}\mu\end{aligned}$$

Example 5.6-1

- Random vector **X** = $(X_1, X_2)^T$ with covariance matrix k = [3 -1;-1 3]. Find transformation **Y=DX** such that **Y**=$(Y_1, Y_2)^T$ is a Normal random vector with uncorrelated (and therefore independent) components of unity variance
- **Solution:** We seek **D** such that $E[\mathbf{YY}^T] = E[\mathbf{DXX}^T\mathbf{D}^T] = \mathbf{DKD}^T = \mathbf{I}$
  - We know that such a transformation is **D = $\Lambda^{-1/2}$U^T**
  - Next, calculate Eigenvalues and Eigenvectors
    - Solving det($\mathbf{K}_{XX}$- $\boldsymbol{\lambda}$**I**) = 0 gets, λ_1 = 4, λ_1 = 2 → $\boldsymbol{\Lambda}^{-1/2} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$
    - Eigenvectors are ($\mathbf{K}_{XX}$ - λ_i **I**) ϕ_i=0,

$$\mathbf{U} = (\phi_1, \phi_2) = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

  - Therefore, we get, $\mathbf{D} = \boldsymbol{\Lambda}^{1/2}\mathbf{U}^T = \frac{1}{\sqrt{2}}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

- Generate correlated vectors **X** from uncorrelated vector **Y** ($\mathbf{K}_{YY}$ is not diagonal) using the transformation **X** = $\mathbf{D}^{-1}$**Y** , where **D** = $\boldsymbol{\Lambda}^{-1/2}\mathbf{U}^T$

Example 5.6-2

X_1 and X_2 are Normal r.v.s with joint pdf $f_{X_1X_2}(x_1,x_2) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho^2}}\exp\left(\frac{-1}{2\sigma^2(1-\rho^2)}(x_1^2 - 2\rho x_1 x_2 + x_2^2)\right)$ and ϱ = -0.5. Find two r.v.'s Y_1 and Y_2 such that they are independent.

- Find K_{XX} in the standard form $\mathbf{x}^T\mathbf{K}_{XX}^{-1}\mathbf{x}$ to diagonalize the cov. Matrix

$$x_1^2 + x_1x_2 + x_2^2 = \mathbf{x}^T\begin{bmatrix} a & b \\ c & d \end{bmatrix}\mathbf{x} = ax_1^2 + (b+c)x_1x_2 + dx_2^2 \longrightarrow \text{a=d = 1 and b=c = 0.5} \longrightarrow \mathbf{K}^{-1} = \frac{1}{\sigma^2(1-\rho^2)}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{4}{3\sigma^2}\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

- Since the constant $4/3\sigma^2$ only affects the eigenvalues and not the eigenvectors. Also we do not need to *whiten* K_{XX} in this problem. Define

$$\tilde{\mathbf{K}}^{-1} \triangleq \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \xrightarrow[\tilde{\lambda}_2 = 1/2]{\tilde{\lambda}_1 = 3/2} \tilde{\mathbf{U}} \triangleq \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \longrightarrow \tilde{\mathbf{U}}^T\tilde{\mathbf{K}}^{-1}\tilde{\mathbf{U}} = \text{diag}(1,3)$$

In Slide 14, we derived $C^TK_{XX}^{-1}C = I$, where $C = UZ^{-1} = U\Lambda^{1/2}$. If we don't want to normalize K_{YY}, then $K_{YY} = U^T[K_{XX}^{-1}]U$ = diag( )

- Hence U^T is a good transformation. $\mathbf{Y} = \mathbf{U}^T\mathbf{X}$ $\longrightarrow$

$$Y_1 = X_1 + X_2$$
$$Y_2 = X_1 - X_2$$

- The pdf is given by $f_{\mathbf{Y}}(\mathbf{y}) = \sum_{i=1}^n f_{\mathbf{X}}(\mathbf{x_i})/|J_i|$

- Using direct method

$$J = \det\left(\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right) = \det\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = -2 \longrightarrow$$

$$f_{Y_1Y_2}(y_1,y_2) = \frac{1}{2}f_{X_1X_2}\left(\frac{y_1+y_2}{2}, \frac{y_1-y_2}{2}\right)$$
$$= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{y_1^2}{2\sigma^2}\right]\cdot\frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left[-\frac{y_2^2}{2\sigma_1^2}\right] \longrightarrow \sigma_1 \triangleq \sqrt{3}\sigma$$

Bivariate Gaussian Using Standard Normal

- Let Z_1 and Z_2 are **standard normal** r.v.
  - We want to find the joint pdf of Y_1,Y_2 with parameters σ_X, σ_Y, μ_X, μ_Y and ρ.
- The transformation for this is $\begin{cases} X = \sigma_X Z_1 + \mu_X \\ Y = \sigma_Y\left[\rho Z_1 + \sqrt{1-\rho^2}Z_2\right] + \mu_Y \end{cases}$
- To check if this is a correct, check marginals are **Gaussian**

$$\begin{aligned} X &= \sigma_X Z_1 + \mu_X \\ &= \sigma_X \mathcal{N}(0,1) + \mu_X \\ &= \mathcal{N}\left(\mu_X, \sigma_X^2\right) \\ Y &= \sigma_Y\left[\rho Z_1 + \sqrt{1-\rho^2}Z_2\right] + \mu_Y \\ &= \sigma_Y[\rho\mathcal{N}(0,1) + \sqrt{1-\rho^2}\mathcal{N}(0,1)] + \mu_Y \\ &= \sigma_Y\left[\mathcal{N}\left(0,\rho^2\right) + \mathcal{N}\left(0, 1-\rho^2\right)\right] + \mu_Y \\ &= \sigma_Y\mathcal{N}(0,1) + \mu_Y \\ &= \mathcal{N}\left(\mu_Y, \sigma_Y^2\right) \end{aligned}$$

Use Y=AX+B transformation.
We know, $f_Y = (1/|a|)f_X[(y-b)/a]$

Use sum of uncorrelated Gaussians

$$\begin{aligned} f_Z(z) &= (f_X * f_Y)(z) \\ &= \mathcal{F}^{-1}\{\mathcal{F}\{f_X\}\cdot\mathcal{F}\{f_Y\}\} \\ &= \mathcal{F}^{-1}\left\{\exp[-j\omega\mu_X]\exp\left[-\frac{\sigma_X^2\omega^2}{2}\right]\exp[-j\omega\mu_Y]\exp\left[-\frac{\sigma_Y^2\omega^2}{2}\right]\right\} \\ &= \mathcal{F}^{-1}\left\{\exp[-j\omega(\mu_X+\mu_Y)]\exp\left[-\frac{(\sigma_X^2+\sigma_Y^2)\omega^2}{2}\right]\right\} \\ &= \mathcal{N}\left(z; \mu_X+\mu_Y, \sigma_X^2+\sigma_Y^2\right) \end{aligned}$$

- Then compute Cov(X,Y) and ρ

$$\begin{aligned} \operatorname{cov}(X,Y) &= E[(X-E(X))(Y-E(Y))] \\ &= E\left[(\sigma_X Z_1 + \mu_X - \mu_X)\left(\sigma_Y\left[\rho Z_1 + \sqrt{1-\rho^2}Z_2\right] + \mu_Y - \mu_Y\right)\right] \\ &= E\left[(\sigma_X Z_1)\left(\sigma_Y\left[\rho Z_1 + \sqrt{1-\rho^2}Z_2\right]\right)\right] \\ &= \sigma_X\sigma_Y E\left[\rho Z_1^2 + \sqrt{1-\rho^2}Z_1Z_2\right] \\ &= \sigma_X\sigma_Y\rho E\left[Z_1^2\right] \\ &= \sigma_X\sigma_Y\rho \end{aligned}$$

$$\rho(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sigma_X\sigma_Y} = \rho$$

So, it is a correct transformation

Joint Density of Bivariate Gaussian

- **(Jacobian Method)** Find are the inverses of the transformation
  - If $X = g(Z_1, Z_2)$ and $Y = h(Z_1, Z_2)$, find functions φ and ψ such that $Z_1 = \varphi(X, Y)$ and $Z_2 = \psi(X, Y)$

$$
\begin{aligned}
X &= \sigma_X Z_1 + \mu_X \\
Z_1 &= \frac{X - \mu_X}{\sigma_X} \\
Y &= \sigma_Y \left[\rho Z_1 + \sqrt{1-\rho^2} Z_2 \right] + \mu_Y \\
\frac{Y - \mu_Y}{\sigma_Y} &= \rho \frac{X - \mu_X}{\sigma_X} + \sqrt{1-\rho^2} Z_2 \\
Z_2 &= \frac{1}{\sqrt{1-\rho^2}} \left[\frac{Y - \mu_Y}{\sigma_Y} - \rho \frac{X - \mu_X}{\sigma_X} \right]
\end{aligned}
$$

Jacobian

$$
\tilde{J} = \det \begin{bmatrix} \frac{\partial \phi}{\partial x} & \frac{\partial \phi}{\partial y} \\ \frac{\partial \psi}{\partial x} & \frac{\partial \psi}{\partial y} \end{bmatrix} = \det \begin{bmatrix} \frac{1}{\sigma_X} & 0 \\ \frac{-\rho}{\sigma_X \sqrt{1-\rho^2}} & \frac{1}{\sigma_Y \sqrt{1-\rho^2}} \end{bmatrix} = \frac{1}{\sigma_X \sigma_Y \sqrt{1-\rho^2}}
$$

- The joint density of X and Y is then given by

$$
\begin{aligned}
f(x, y) &= f(z_1, z_2)|J| \\
&= \frac{1}{2\pi} \exp\left[-\frac{1}{2} \left(z_1^2 + z_2^2 \right) \right] |\jmath = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} \exp\left[-\frac{1}{2} \left(z_1^2 + z_2^2 \right) \right] \\
&= \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1-\rho^2}} \exp\left[-\frac{1}{2} \left[\left(\frac{x - \mu_X}{\sigma_X} \right)^2 + \frac{1}{1-\rho^2} \left(\frac{y - \mu_Y}{\sigma_Y} - \rho \frac{x - \mu_X}{\sigma_X} \right)^2 \right] \right] \\
&= \frac{1}{2\pi \sigma_X \sigma_Y \left(1-\rho^2\right)^{1/2}} \exp\left[\frac{-1}{2\left(1-\rho^2\right)} \left(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_X^2} - 2\rho \frac{(x-\mu_X)}{\sigma_X} \frac{(y-\mu_Y)}{\sigma_Y} \right) \right]
\end{aligned}
$$

$(\det \mathbf{K})^{1/2}$

$(x-\mu)^T \mathbf{K}^{-1} (x-\mu)$

Compare with the Matrix Notation

- From the previous slides we write in matrix form

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \boldsymbol{\mu} = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \quad \mathbf{K} = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix} \longrightarrow f(\mathbf{x}) = \frac{1}{2\pi(\det\mathbf{K})^{-1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{K}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$$

- Check if $(\det \mathbf{K})^{1/2}$ and $(x\text{-}\mu)^T \mathbf{K}^{-1} (x\text{-}\mu)$ matches the regular form above

$$\longrightarrow \quad (\det\mathbf{K})^{-1/2} = \left(\sigma_X^2\sigma_Y^2 - \rho^2\sigma_X^2\sigma_Y^2\right)^{-1/2} = \frac{1}{\sigma_X\sigma_Y\left(1-\rho^2\right)^{1/2}}$$

- Also recall $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \mathbf{A}^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$
- Then the exponent of the bivariate Gaussian is given by

$$\begin{aligned}
&(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \\
&= \frac{1}{\sigma_X^2\sigma_Y^2(1-\rho^2)}\begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix}^T \begin{pmatrix} \sigma_Y^2 & -\rho\sigma_X\sigma_Y \\ -\rho\sigma_X\sigma_Y & \sigma_X^2 \end{pmatrix}\begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix} \\
&= \frac{1}{\sigma_X^2\sigma_Y^2(1-\rho^2)}\begin{pmatrix} \sigma_Y^2(x-\mu_X) - \rho\sigma_X\sigma_Y(y-\mu_Y) \\ -\rho\sigma_X\sigma_Y(x-\mu_X) + \sigma_X^2(y-\mu_Y) \end{pmatrix}^T \begin{pmatrix} x-\mu_x \\ y-\mu_y \end{pmatrix} \\
&= \frac{1}{\sigma_X^2\sigma_Y^2(1-\rho^2)}\left(\sigma_Y^2(x-\mu_X)^2 - 2\rho\sigma_X\sigma_Y(x-\mu_X)(y-\mu_Y) + \sigma_X^2(y-\mu_Y)^2\right) \\
\longrightarrow \; &= \frac{1}{1-\rho^2}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} - 2\rho\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right)
\end{aligned}$$

Characteristic Function

- Similar to scalar r.v. we have the c.f. for **X**=[X_1, X_2, ... ,X_N]^T

$$\Phi_{\mathbf{X}}(\omega) \triangleq E\left[e^{j\omega^T\mathbf{X}}\right] = \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) e^{j\omega^T\mathbf{x}} d\mathbf{x} \quad \xleftrightarrow{\text{Inverse}} \quad f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty} \Phi_{\mathbf{X}}(\omega) e^{-j\omega^T\mathbf{x}} d\omega$$

- Similar to scalar r.v. moments (if they exist) can be found using the c.f.
  - Example 5.7-1: For **X**=[X_1, X_2,X_3]^T and $\boldsymbol{\omega}$ = [ω_1,ω_2,ω_3], we write

$$\Phi_{\mathbf{X}}(\omega_1,\omega_2,\omega_3) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbf{x}}(x_1,x_2,x_3) e^{j[\omega_1 x_1+\omega_2 x_2+\omega_3 x_3]} dx_1 dx_2 dx_3$$

  - By partially deriving

$$\frac{1}{j^3}\frac{\partial^3 \Phi_{\mathbf{X}}(\omega_1,\omega_2,\omega_3)}{\partial\omega_1\partial\omega_2\partial\omega_3}\bigg|_{\omega_1=\omega_2=\omega_3=0} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2x_3 f_{\mathbf{X}}(x_1,x_2,x_3)dx_1dx_2dx_3 \triangleq E[X_1X_2X_3]$$

- Therefore, generalizing the above expression to n-dimensions,

$$E\left[X_1^{k_1}\dots X_n^{k_n}\right] = j^{(k_1+\dots+k_n)}\frac{\partial^{k_1+\dots+k_n}\Phi_{\mathbf{X}}(\omega_1,\dots,\omega_n)}{\partial\omega_1^{k_1}\dots\partial\omega_n^{k_n}}\bigg|_{\omega_1=\dots=\omega_n=0}$$

- We can also write the c.f. as products of exponentials

$$E[\exp(j\omega^T\mathbf{X})] = E\left[\exp\left(j\sum_{i=1}^{n}\omega_i X_i\right)\right] = E\left[\prod_{i=1}^{n}\exp(j\omega_i X_i)\right]$$

Recall for two variables, the joint moments are a below

$$m_{rk} \triangleq E\left[X^rY^k\right] = (-j)^{r+k}\Phi_{XY}^{(r,k)}(0,0)$$

$$\Phi_{XY}^{(rk)}(0,0) \triangleq \frac{\partial^{r+k}\Phi_{XY}(\omega_1,\omega_2)}{\partial\omega_1^r\partial\omega_2^k}\bigg|_{\omega_1=\omega_2=0}$$

Properties of CF

- Properties :

1. $|\Phi_{\mathrm{X}}(\omega)| \leq \Phi_{\mathrm{X}}(0) = 1$ and
2. $\Phi^*_{\mathrm{X}}(\omega) = \Phi_{\mathrm{X}}(-\omega)$($*$ indicates conjugation).
3. All c.f.'s of subsets of the components of X can be obtained once $\Phi_{\mathrm{X}}(\omega)$ is known.

- An example of the last property is

$$\begin{aligned}\Phi_{X_1X_2}(\omega_1,\omega_2) &= \Phi_{X_1X_2X_3}(\omega_1,\omega_2,0)\\ \Phi_{X_1X_3}(\omega_1,\omega_3) &= \Phi_{X_1X_2X_3}(\omega_1,0,\omega_3)\\ \Phi_{X_1}(\omega_1) &= \Phi_{X_1X_2X_3}(\omega_1,0,0)\end{aligned}$$

- Best application of c.f.: Convert convolution of r.v. to multiplication of c.f.

$$\begin{gathered} Z = X_1 + \ldots + X_n \\ f_Z(z) = f_{X_1}(z) * \ldots * f_{X_n}(z) \end{gathered}$$

Using CF →

$$\begin{aligned}\Phi_{\mathbf{X}}(\mathbf{z}) &= E\left[e^{j\omega(X_1+\ldots+X_n)}\right]\\ &= E\left[\prod_{i=1}^{n} e^{j\omega X_i}\right] = \prod_{i=1}^{n} E\left[e^{j\omega X_i}\right] = \prod_{i=1}^{n} \Phi_{X_i}(\omega)\end{aligned}$$

Example 5.7-2

- If X_i are iid Poisson random variable X_i. Find the pdf of Z = $X_1 + X_2 + ... + X_N$?
- Characteristic function of a Poisson r.v. X is given by

$$\begin{aligned}\Phi_X(\omega) &= E\left[e^{j\omega X}\right] = \sum_{k=0}^{\infty} e^{j\omega k} P_X(k) \\ &= \sum_{k=0}^{\infty} e^{j\omega k} \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^{j\omega}\right)^k}{k!} \\ &= e^{-\lambda} e^{\lambda e^{i\omega}} = e^{\lambda(e^{i\omega}-1)}\end{aligned}$$

- The pdf of Z is given by

$$\begin{aligned}\Phi_Z(\omega) &= \prod_{i=1}^{n} e^{\lambda(e^{j\omega}-1)} \\ &= e^{n\lambda(e^{j\omega}-1)}\end{aligned}$$

- Which is a CF of a Poisson r.v. with parameter $\alpha = n\lambda$, i.e.

$$P_Z(k) = \frac{\alpha^k e^{-\alpha}}{k!}$$

Characteristic Function of Gaussian

- Let **X** be a normal random vector with nonsingular covariance matrix **K**, then both **K** and **K**$^{-1}$ can be factored as $\mathbf{K} = \mathbf{C}\mathbf{C}^T \longrightarrow \mathbf{K}^{-1} = \mathbf{D}\mathbf{D}^T, \quad \mathbf{D} = [\mathbf{C}^T]^{-1}$
- The CF for the normal r.v. is

$$\Phi_{\mathbf{X}}(\omega) = \frac{1}{(2\pi)^{n/2}[\det(\mathbf{K})]^{1/2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{K}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) \cdot \exp\left(j\boldsymbol{\omega}^T\mathbf{x}\right) d\mathbf{x}$$

- Under the transformation $\mathbf{z} \triangleq \mathbf{D}^T(\mathbf{x}-\mu) \implies \mathbf{x} = [\mathbf{D}^T]^{-1}\mathbf{z} + \mu$ we get

$$\begin{aligned} \mathbf{z}^T\mathbf{z} &= (\mathbf{x}-\boldsymbol{\mu})^T\mathbf{D}\mathbf{D}^T(\mathbf{x}-\boldsymbol{\mu}) \\ &= (\mathbf{x}-\boldsymbol{\mu})^T\mathbf{K}^{-1}(\mathbf{x}-\boldsymbol{\mu}) \end{aligned} \qquad d\mathbf{z} = |\det(\mathbf{D}^T)|d\mathbf{x}$$

Therefore, the CF after transformation

$$\Phi_{\mathbf{X}}(\omega) = \frac{\exp\left(j\omega^T\mu\right)}{(2\pi)^{n/2}[\det(\mathbf{K})]^{1/2}|\det(\mathbf{D})|} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\mathbf{z}^T\mathbf{z}\right) \cdot \exp\left(j\omega^T\left(\mathbf{D}^T\right)^{-1}\mathbf{z}\right) d\mathbf{z}$$

Swap inverse and transpose, since $(A^T)^{-1} = (A^{-1})^T$

- Complete the square in the integrand

$$\exp\left[-\frac{1}{2}\left[\mathbf{z}^T\mathbf{z} - 2j\boldsymbol{\omega}^T\left(\mathbf{D}^T\right)^{-1}\mathbf{z}\right]\right] = \exp\left[-\frac{1}{2}\left[\mathbf{z}^T\mathbf{z} - 2j\left(\boldsymbol{\omega}^T\left(\mathbf{D}^{-1}\right)^T\right)\mathbf{z}\right]\right] = \exp\left[-\frac{1}{2}\left[\mathbf{z}^T\mathbf{z} - 2j\left(\mathbf{D}^{-1}\boldsymbol{\omega}\right)^T\mathbf{z}\right]\right]$$

$$= \exp\left(-\frac{1}{2}\boldsymbol{\omega}^T\left(\mathbf{D}^T\right)^{-1}\mathbf{D}^{-1}\boldsymbol{\omega}\right) \cdot \exp\left(-\frac{1}{2}\left\|\mathbf{z} - j\mathbf{D}^{-1}\boldsymbol{\omega}\right\|^2\right)$$

$$x^TMx - 2b^Tx = \left(x - M^{-1}b\right)^T M\left(x - M^{-1}b\right) - b^TM^{-1}b$$

$$(\mathbf{D}\mathbf{D}^T)^{-1} = \mathbf{K} \quad \text{and} \quad \det(\mathbf{K}^{-1}) = [\det(\mathbf{K})]^{-1} = \det(\mathbf{D})\det(\mathbf{D}^T) = \det[(\mathbf{D})]^2 \implies |\det(\mathbf{D}| = \det(\mathbf{K})^{-1/2}$$

... contd

- So, the CF reduces to

$$\Phi_{\mathbf{X}}(\omega) = \exp\left(j\omega^T\mu - \frac{1}{2}\omega^T\mathbf{K}\omega\right) \cdot \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left\|\mathbf{z} - j\mathbf{D}^{-1}\omega\right\|^2} d\mathbf{z}$$

- The integral is a n-fold integration of n iid rv with unit variance = $(2\pi)^{n/2}$
- Hence the cf of a Gaussian maps onto a Gaussian cf

$$\Phi_{\mathbf{X}}(\omega) = \exp\left(j\omega^T\mu - \frac{1}{2}\omega^T\mathbf{K}\omega\right)$$