Lecture 5

ECE 278 Mathematics for MS Comp Exam

- A **quadratic form** Q of the vector $\mathbf{x}$ with real components x_i is given by the matrix representation $\mathbb{A}$ of that transformation that satisfies

$$Q = \mathbf{x}^T \mathbb{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k \tag{1}$$

- The matrix $\mathbb{A}$ is called the coefficient matrix
  - learn in probability that a **covariance matrix** is an example of a coefficient matrix
- Sum leads to n “diagonal terms” when $j = k$ and $n^2 - n$ “off-diagonal terms ” when $j \neq k$
- For a vector $\mathbf{x}$ with complex components

$$Q = \mathbf{x}^H \mathbb{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j^* x_k$$

- From previous lecture, when $\mathbb{A}$ is symmetric (hermitian) it has an orthonormal set of eigenvectors
- Forming a matrix from these the unit norm column eigenvectors gives a matrix $\mathbb{X}$ that is **orthogonal (unitary)**
  - For an orthogonal (unitary) matrix $\mathbb{X}^{-1} = \mathbb{X}^T$ (or $\mathbb{X}^{-1} = \mathbb{X}^H$)
- The transformation be viewed as a rotation (or a generalized rotation) that preserves the **norm** (length)
- The generalized rotation matrix $\mathbb{X}$ can be used to diagonalize $\mathbb{A}$ as follows

$$\mathbb{A} = \mathbb{X}\mathbb{D}\mathbb{X}^{-1} = \mathbb{X}\mathbb{D}\mathbb{X}^T \tag{2}$$

where the diagonal elements are the eigenvalues λ_n

- Substitute (2) into (1) to give

$$\begin{aligned} Q &= \mathbf{x}^T \mathbb{X}\mathbb{D}\mathbb{X}^T \mathbf{x} \\ &= \mathbf{y}^T \mathbb{D}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 \end{aligned} \tag{3}$$

  where

$$\mathbf{y} = \mathbb{X}^T \mathbf{x} \qquad \mathbf{y}^T = \mathbf{x}^T \mathbb{X}$$

- Because $\mathbb{X}$ is orthogonal, $\mathbb{X}^T = \mathbb{X}^{-1}$. Substituting into first equation gives

$$\mathbf{y} = \mathbb{X}^{-1}\mathbf{x}$$

  or

$$\boxed{\mathbf{x} = \mathbb{X}\mathbf{y}}$$

- Expression states that by performing a **change of basis** expressed by the matrix $\mathbb{X}$(from the "new" coordinate system $\mathbf{y}$ to the "old" coordinate system $\mathbf{x}$)

- The change of basis transforms the original quadratic form (with both on-diagonal and off-diagonal terms) can be transformed into a **canonical quadratic form** that has only "on-diagonal terms" as given by (3).

- Problem Set 8.4 #17.

$$7x^2 + 6xy + 7y^2 - 200 = C$$

(This is a family of ellipses instead of a single ellipse)

- Rewrite as

$$7x^2 + 3xy + 3yx + 7y^2 - 200 = C$$

- Coefficient matrix is then

$$\mathbb{A} = \begin{bmatrix} 7 & 3 \\ 3 & 7 \end{bmatrix}$$

- Eigenvalues are solution to $(7-\lambda)^2 - 9 = 0$ and are $\lambda_1 = 4$ and $\lambda_2 = 10$

- Because matrix is symmetric, normalized eigenvectors form an orthonormal basis given by

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

- Form the matrix that describes the change of basis using eigenvectors

$$\mathbb{X} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad \mathbb{X}^{-1} = \mathbb{X}^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

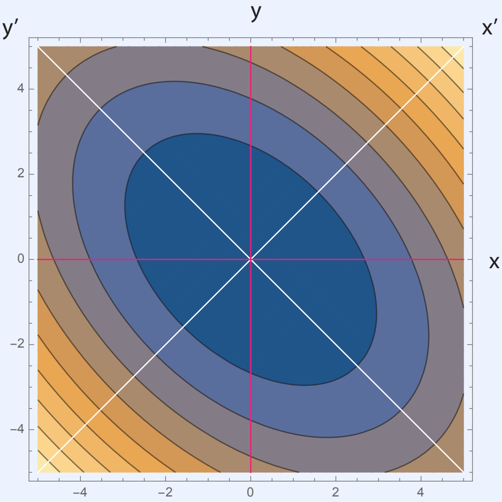
- So that

$$\mathbb{D} = \mathbb{X}^T \mathbb{A} \mathbb{X} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 7 & 3 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 10 \end{bmatrix}$$

- New $\mathbf{x}'$ axis can be determined using change of basis transformation as given by

$$\mathbf{x}' = \mathbb{X}^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathbb{X}^T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- Therefore $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \longrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. This is a rotation of 45^o


y
y′
x′
x
4
2
0
−2
−4
−4
−2
0
2
4

- A ***multivariate gaussian probability density function*** is a joint probability density function for a block $\mathbf{x}$ of real random variables with components $\underline{x}_i$ given by

$$f_{\underline{\mathbf{x}}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^N \det \mathbb{C}}} e^{-\frac{1}{2}(\mathbf{x}-\langle\mathbf{x}\rangle)^T \mathbb{C}^{-1}(\mathbf{x}-\langle\mathbf{x}\rangle)}, \tag{4}$$

- Argument of gaussian is a **quadratic form**
- The coefficient matrix $\mathbb{C}$ is the ***real covariance matrix*** defined for any multivariate probability density function as

$$\mathbb{C} = \left\langle \left(\underline{\mathbf{x}} - \langle\mathbf{x}\rangle\right)\left(\underline{\mathbf{x}} - \langle\mathbf{x}\rangle\right)^T \right\rangle. \tag{5}$$

- The square symmetric matrix $\mathbb{C}$ has a determinant $\det \mathbb{C}$.
- The diagonal matrix element C_{ii} is the variance of the random variable $\underline{x}_i$.
- The off-diagonal matrix element C_{ij} is the covariance of the two random variables $\underline{x}_i$ and $\underline{x}_j$.

- A zero-mean bivariate gaussian random variable consists of two random, zero-mean gaussian components $\underline{x}$ and $\underline{y}$, which may be correlated.
- The covariance matrix given in (5) is

$$\mathbb{C} = \begin{bmatrix} \sigma_x^2 & \rho_{xy}\sigma_x\sigma_y \\ \rho_{xy}\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}, \tag{6}$$

where

$$\rho_{xy} \doteq \left\langle \underline{xy} \right\rangle / \sigma_x\sigma_y, \tag{7}$$

is defined as the ***correlation coefficient***.

- Describes value of the off-diagonal elements of covariance matrix.

An example of a two-dimensional joint gaussian probability density function is shown in plan view in Figure 1.

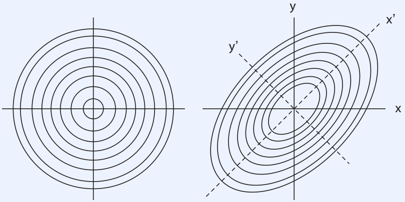


Figure: Contours of the joint gaussian probability density function $f_{\underline{x},\underline{y}}(x, y)$ as a function of the correlation coefficient ρ_{xy}: (a) $\rho_{xy} = 0$, (b) $\rho_{xy} = 0.5$.

- If $\sigma_x = \sigma_y = \sigma$, then (4) reduces to

$$f_{\underline{x},\underline{y}}(x,y) = \frac{1}{2\pi\sigma^2\sqrt{1-\rho_{xy}^2}} \exp\left(-\frac{x^2 - 2\rho_{xy}xy + y^2}{2\sigma^2(1-\rho_{xy}^2)}\right). \tag{8}$$

UCSD Jacobs | Electrical and Computer Engineering

- The set of all complex signals of finite energy on an interval $[0, T]$ defines the *signal space* over that interval.
- Two elements within the signal space are deemed to be equivalent or equal, if the energy of the difference of the two signals is zero.
- A countable set of signals $\{\psi_n(t)\}$ of a signal space *spans* that signal space if every element of the signal space can be expressed as a linear combination of the $\psi_n(t)$.
- This means that every $s(t)$ can be written as

$$s(t) \quad = \quad \sum_n s_n \psi_n(t), \tag{9}$$

in the sense that the difference between the left side and the right side has zero energy.

UCSD Jacobs | Electrical and Computer Engineering

- The set $\{\psi_n(t)\}$ is called a ***basis*** if the elements of the set are linearly independent and span the signal space. Every basis for a signal space is countably infinite.

- An **orthonormal basis** $\{\psi_n(t)\}$ satisfies the additional requirement that

$$\int_0^T \psi_m(t)\psi_n^*(t)\mathrm{d}t \quad \doteq \quad \delta_{mn}, \tag{10}$$

  for all m and n, where δ_{mn} is the Kronecker impulse

- This means that each basis function satisfies

$$\int_0^T |\psi_n(t)|^2\mathrm{d}t \quad = \quad 1. \tag{11}$$

  and

$$\int_0^T \psi_m(t)\psi_n^*(t)\mathrm{d}t \quad = \quad 0 \qquad (m \neq n). \tag{12}$$

- For an orthonormal basis, the coefficient s_n of the expansion in (9) is given by

$$s_n \doteq \int_0^T s(t)\psi_n^*(t)\mathrm{d}t. \qquad (13)$$

- A set of basis functions must span the entire signal space,
  - implies that the number of functions in any basis for signal space is infinite.
  - For a finite time internal (like Fourer series) this is a countable infinity - can be mapped onto the integers
  - For and infinite time interval (like a Fourier transform) this is an uncountable infinity
- A basis must be infinite, but not every infinite orthonormal set is a basis for the set of square-integrable functions.
- A linear transformation on a signal space is a mapping from the space onto itself that satisfies the linearity properties.
- With respect to a fixed basis, a linear transformation can be described by a matrix, called the ***transformation matrix***.

- For any orthonormal basis $\{\psi_m(t)\}$ a signal $s(t)$ in the signal space over $[0, T]$ is completely determined by an infinite sequence of complex components s_n, which are the coefficients of the expansion given in (9).
- These coefficients may be regarded as forming an infinitely long vector s called a ***signal vector***.
- Given a signal vector s with complex components, define the *conjugate transpose vector* as the vector $\mathsf{s}^\dagger$ whose components are the complex conjugates of the corresponding components of the vector s. If s is a column vector, then $\mathsf{s}^\dagger$ is a row vector.

- Using $a(t) = \sum_m a_m \psi_m(t)$ (cf. (9)) and $b(t) = \sum_n b_n \psi_n(t)$, define the ***inner product*** as

$$
\begin{aligned}
\mathbf{a} \cdot \mathbf{b} &\doteq \int_0^T a(t) b^*(t) \mathrm{d}t \\
&= \sum_m \sum_n a_m b_n^* \int_0^T \psi_m(t) \psi_n^*(t) \mathrm{d}t \\
&= \sum_m \sum_n a_m b_n^* \delta_{mn} \\
&= \sum_m a_m b_m^*, \qquad (14)
\end{aligned}
$$

  where (10) is used in going from the second line to the third line.
- Setting $a(t) = b(t)$ in (14) immediately gives the energy statement

$$
\int_0^T |a(t)|^2 \mathrm{d}t = \sum_m |a_m|^2. \qquad (15)
$$

- For a finite-energy signal $a(t)$, this implies that $|a_m|^2$ goes to zero as m goes to infinity.
- For some integer N, an arbitrarily small amount of energy is discarded by including only N terms.

- The term a_m is a component of the (column) signal vector $\mathbf{a}$. The term b_m^* is the component of a (row) signal vector $\mathbf{b}^\dagger$. Therefore,

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b}^\dagger \mathbf{a}, \tag{16}$$

where $\mathbf{b}^\dagger \mathbf{a}$ is the matrix product of a one by N matrix and an N by one matrix.

- Using (15), the energy in the signal $s(t)$ is

$$E = \int_0^T |s(t)|^2 \mathrm{d}t = \sum_n |s_n|^2 = |\mathbf{s}|^2. \tag{17}$$

- Similarly, the component s_n in (13) is determined using $a(t) = s(t)$ and $b(t) = \psi_n(t)$

$$s_n = \int_0^T s(t)\psi_n^*(t)\mathrm{d}t \doteq \mathbf{s} \cdot \boldsymbol{\psi}_n. \tag{18}$$

- This expression defines the *projection* of $s(t)$ onto $\psi_n(t)$.
- The vector $\boldsymbol{\psi}_n$ has the nth component equal to one and all other components equal to zero.
- It is a *basis vector* that corresponds to the basis function $\psi_n(t)$ defined in (9).

- Using (14) and (17), the *squared euclidean distance* d_{12}^2 between two signals $s_1(t)$ and $s_2(t)$, or, equivalently two signal vectors $\mathbf{s}_1$ and $\mathbf{s}_2$, is defined as

$$\begin{aligned} d_{12}^2 &\doteq \int_0^T |s_1(t) - s_2(t)|^2 \, \mathrm{d}t \\ &= \int_0^T \Big(|s_1(t)|^2 + |s_2(t)|^2 - s_1(t)s_2^*(t) - s_1^*(t)s_2(t) \Big) \mathrm{d}t \\ &= E_1 + E_2 - 2\mathrm{Re}\left[\mathbf{s}_1 \cdot \mathbf{s}_2\right]. \end{aligned} \tag{19}$$

- This expression states that the squared euclidean distance between two signals depends on the energy of each signal as well as their inner product.