Stochastic Processes

Topic 3

More on Joint Gaussian Density

`nctuee09f`

Summary

In this lecture, I will discuss:

- Conditional Jointly Gaussian Density
- Complex Gaussian Random Vector
- A Simple Detection Problem

Notation

We will use the following notation rules, unless otherwise noted, to represent symbols during this course.

- Boldface upper case letter to represent **MATRIX**
- Boldface lower case letter to represent **vector**
- Superscript $(\cdot)^T$ and $(\cdot)^H$ to denote transpose and hermitian (conjugate transpose), respectively
- Upper case italic letter to represent *RANDOM VARIABLE*

1 Conditional Joint Gaussian Density

Consider a jointly Gaussian random vector $[\mathbf{x}^T\ \mathbf{y}^T]^T$ with $\mathbf{x} \sim \mathcal{N}(\mathbf{m_x}, \mathbf{K_x})$ and $\mathbf{y} \sim \mathcal{N}(\mathbf{m_y}, \mathbf{K_y})$. Then, the conditional density $f_{\mathbf{x}|\mathbf{y}}(\boldsymbol{x}|\boldsymbol{y})$ also follows a joint Gaussian density with

$$\mathbf{x}|\mathbf{y} \sim \mathcal{N}\Bigg(\ \mathbf{m_x} + \mathbf{K_{xy}}\mathbf{K_y}^{-1}(\mathbf{y} - \mathbf{m_y}),\ \mathbf{K_x} - \mathbf{K_{xy}}\mathbf{K_y}^{-1}\mathbf{K_{yx}}\Bigg),$$

where $\mathbf{K_{xy}} = E\left[(\mathbf{x} - \mathbf{m_x})(\mathbf{y} - \mathbf{m_y})^H\right]$ and $\mathbf{K_{yx}} = \mathbf{K}_{\mathbf{xy}}^H$

Proof

(1) Using $f_{\mathbf{x}|\mathbf{y}}(\boldsymbol{x}|\boldsymbol{y}) = \frac{f_{\mathbf{x},\mathbf{y}}(\boldsymbol{x},\boldsymbol{y})}{f_{\mathbf{y}}(\boldsymbol{y})}$.
(Tedious process and not obvious. See Gallager's note.)

(2) As an alternative approach, we first find the matrix $\mathbf{A}$ such that the random vector

$$\mathbf{z} = (\mathbf{x} - \mathbf{m_x}) - \mathbf{A}(\mathbf{y} - \mathbf{m_y})$$

is independent with $\mathbf{y}$. It is clear that $\mathbf{z}$ is a linear combination of $\mathbf{x}$ and $\mathbf{y}$, and therefore is also a ***Gaussian random vector*** (*i.e.* all components of $\mathbf{z}$ are jointly Gaussian) with mean vector zero and covariance matrix $\mathbf{K_x} - \mathbf{K_{xy}}\mathbf{K}_{\mathbf{y}}^{-1}\mathbf{K_{yx}}$.

The idea here is to express $\mathbf{x}$ by two terms, one independent with $\mathbf{y}$ and the other solely dependent on $\mathbf{y}$. That is

$$\mathbf{x} = \mathbf{z} + \mathbf{x} - \mathbf{z}$$

By doing so, the conditional random vector $\mathbf{x}|\mathbf{y}$ can also be expressed in two terms, one is constant due to the conditioning on $\mathbf{y}$ and the other, i.e. the random vector $\mathbf{z}$, is the only random term contained in $\mathbf{x}|\mathbf{y}$. Hence, we know $\mathbf{x}|\mathbf{y}$ is also a Gaussian random vector with mean vector

$$\begin{aligned} E[\mathbf{x}|\mathbf{y}] &= E\Big[\mathbf{z}\Big|\mathbf{y}\Big] + E\Big[\Big(\mathbf{x} - \mathbf{z}\Big)\Big|\mathbf{y}\Big] \\ &= E\Big[\mathbf{z}\Big] + E\Big[\Big(\mathbf{m_x} + \mathbf{A}(\mathbf{y} - \mathbf{m_y})\Big)\Big|\mathbf{y}\Big] \\ &= \mathbf{m_x} + \mathbf{K_{xy}}\mathbf{K_y}^{-1}(\mathbf{y} - \mathbf{m_y}), \end{aligned}$$

and covariance matrix

$$\text{Cov}(\mathbf{x}|\mathbf{y}) = \text{Cov}(\mathbf{z}) = \mathbf{K_x} - \mathbf{K_{xy}}\mathbf{K}_{\mathbf{y}}^{-1}\mathbf{K_{yx}}.$$

Remarks:

— The conditional mean vector $E[\mathbf{x}|\mathbf{y}]$ is also a Gaussian random vector (all components are jointly Gaussian) depending on $\mathbf{y}$ only.

— The variance of each element in $\mathbf{x}$ is reduced due to the observation of $\mathbf{y}$.

$\Longrightarrow$ Observation of $\mathbf{y}$ gives us additional information about $\mathbf{x}$, thus reducing the variance of it.

— It will be shown in later lectures that the ***minimum mean-square error (MMSE)*** estimate of $\mathbf{x}$ based on the observation $\mathbf{y}$ is given by

$$\hat{\mathbf{x}}_{MMSE} = E[\mathbf{x}|\mathbf{y}],$$

which is generally a nonlinear function of $\mathbf{y}$.

When $\mathbf{x}$ and $\mathbf{y}$ are jointly Gaussian, the MMSE estimate is ***linear*** in $\mathbf{y}$.

2 Complex Gaussian Vector

(1) A ***complex Gaussian*** vector $\mathbf{x} = \mathbf{x}_r + j\mathbf{x}_i$ is a random vector whose real part $\mathbf{x}_r$ and imaginary part $\mathbf{x}_i$ are collectively jointly Gaussian.

— The joint pdf of an n-dimensional complex random vector $\mathbf{x} = \mathbf{x}_r + j\mathbf{x}_i$ is ***defined*** to be the joint pdf of the $2n$-dimensional real random vector $[\mathbf{x}_r^T, \mathbf{x}_i^T]^T$.

— A complex Gaussian vector is completely specified by the mean $\mathbf{m}_\mathbf{x} = E[\mathbf{x}]$, the covariance matrix

$$\mathbf{K}_\mathbf{x} = E[(\mathbf{x} - \mathbf{m}_\mathbf{x})(\mathbf{x} - \mathbf{m}_\mathbf{x})^H],$$

and the *pseudo-covariance* matrix

$$\mathbf{J}_\mathbf{x} = E[(\mathbf{x} - \mathbf{m}_\mathbf{x})(\mathbf{x} - \mathbf{m}_\mathbf{x})^T]$$

of the complex vector $\mathbf{x}$.

(2) In applications of wireless communications, we are almost exclusively interested in complex random vectors that have the ***circular symmetry*** property:

$\mathbf{x}$ is ***circularly symmetric*** if $e^{j\theta}\mathbf{x}$ has the same distribution as $\mathbf{x}$ for any θ

— For a circularly symmetric complex random vector $\mathbf{x}$, its mean vector is a zero vector and its pseudo-covariance matrix is a zero matrix.

— A ***circularly symmetric complex Gaussian*** random vector $\mathbf{x}$ is completely specified by its covariance matrix and is denoted as $\mathbf{x} \sim \mathcal{CN}(0, \mathbf{K}_\mathbf{x})$.

(3) A complex Gaussian random variable $X = X_r + jX_i$ with i.i.d. zero mean Gaussian real and imaginary components is circular symmetric.

— The statistics of X are fully specified by the variance $\sigma^2 = E[|X|^2]$, and is denoted as $X \sim \mathcal{CN}(0, \sigma^2)$.

— We can represent X in polar form $X = ||X||e^{j\Theta}$, where the phase Θ is uniform over $[0, 2\pi]$ and independent of the magnitude $||X|| = (X_r^2 + X_i^2)^{1/2}$, which has a density given by

$$f_{||X||}(r) = \frac{r}{\sigma^2} \exp\left(\frac{-r^2}{2\sigma^2}\right), \quad r \geq 0$$

and is known as a ***Rayleigh*** random variable.

See Example 3.3-9 (p. 150) in textbook for a derivation of the Rayleigh density, and Example 3.3-10 (p. 151) for the Rician density.

(4) A collection of n i.i.d. $\mathcal{CN}(0,1)$ random variables forms a standard circularly symmetric Gaussian random vector $\mathbf{w}$ and is denoted by $\mathcal{CN}(0,\mathbf{I})$. The joint density of $\mathbf{w}$ is given by

$$f_{\mathbf{w}}(\mathsf{w}) = \frac{1}{\pi^n} \exp\left(-||\mathsf{w}||^2\right).$$

(5) If $\mathbf{x} \sim \mathcal{CN}(0,\mathbf{K_x})$ and $\mathbf{K_x}$ is invertible (non-singular), then the joint density of $\mathbf{x}$ is

$$\boxed{f_{\mathbf{x}}(\mathsf{x}) = \tfrac{1}{\pi^n \det(\mathbf{K_x})} \exp\left(-\mathsf{x}^H \mathbf{K_x}^{-1}\mathsf{x}\right)}$$

where $\mathbf{K_x} = E[(\mathbf{x}-\mathbf{m_x})(\mathbf{x}-\mathbf{m_x})^H]$ is the covariance matrix of $\mathbf{x}$ [1, App. A], [2, Chap 4].

3 A Simple Detection Problem

Suppose we want to transmit binary data thorough a communication link, where we use $\mathbf{s}_0$ and $\mathbf{s}_1$ to represent 0 and 1, respectively. The receiver receives

$$\mathbf{x} = \mathbf{s} + \mathbf{n},$$

where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ is the additive white Gaussian noise. The task of the receiver is to decide between the following two hypotheses:

$$\begin{aligned} \mathtt{H}_0: &\quad \mathbf{x} = \mathbf{s}_0 + \mathbf{n} \\ \mathtt{H}_1: &\quad \mathbf{x} = \mathbf{s}_1 + \mathbf{n}. \end{aligned}$$

The maximum likelihood principle gives the following rule:

$$\boxed{f_{\mathbf{x}}(\mathsf{x}|\mathbf{s}_0) \gtrless_{\mathtt{H}_1}^{\mathtt{H}_0} f_{\mathbf{x}}(\mathsf{x}|\mathbf{s}_1),}$$

where $f_{\mathbf{x}}(\mathsf{x}|\mathbf{s}_0)$ and $f_{\mathbf{x}}(\mathsf{x}|\mathbf{s}_1)$ are the likelihood functions of $\mathbf{x}$ associated with $\mathbf{s}_0$ and $\mathbf{s}_1$, respectively.

Carrying out the above likelihood function, the decision rule is

$$\exp\left(-\frac{1}{2\sigma^2}(\mathsf{x}-\mathbf{s}_0)^T(\mathsf{x}-\mathbf{s}_0)\right) \underset{\mathtt{H}_1}{\overset{\mathtt{H}_0}{\gtrless}} \exp\left(-\frac{1}{2\sigma^2}(\mathsf{x}-\mathbf{s}_1)^T(\mathsf{x}-\mathbf{s}_1)\right).$$

With some straightforward manipulations, we have

$$\left|\left|\mathsf{x}-\mathbf{s}_0\right|\right| \underset{\mathtt{H}_0}{\overset{\mathtt{H}_1}{\gtrless}} \left|\left|\mathsf{x}-\mathbf{s}_1\right|\right|,$$

which is the so called "***distance rule***." With further algebraic efforts, it can be obtained that the receiver actually needs to conduct the following operation

$$\mathsf{x}^T\left(\mathbf{s}_0-\mathbf{s}_1\right) \gtrless^{\mathrm{H}_0}_{\mathrm{H}_1} \frac{1}{2}\left(\left|\left|\mathbf{s}_0\right|\right|^2-\left|\left|\mathbf{s}_1\right|\right|^2\right)$$

to decide between H_0 and H_1.

What is the probability of error decision?

References

[1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge, 2005.

[2] Robert Gallager, *Gallager's Notes on Random Process.*