eMasters in Communication Systems

Prof. Aditya Jagannatham

Elective Module:

Detection for Wireless Communication

Chapter 6

Linear Discriminant Analysis (LDA)

Gaussian Discriminant Analysis (GDA)

Gaussian Classification

n dimensional Random Vector

$R^{n\times 1}$

- Recall, the PDF of a **Gaussian** ***random vector*** is given as

$$E\{\bar{x}\} = \bar{\mu}$$

covariance = R

$$E\{(\bar{x}-\bar{\mu})(\bar{x}-\bar{\mu})^T\} = R$$

$$\left.\frac{1}{\sqrt{(2\pi)^n |R|}} \cdot e^{-\frac{1}{2}(\bar{x}-\bar{\mu})^T R^{-1}(\bar{x}-\bar{\mu})}\right\}$$

Multivariate Gaussian PDF.

$\bar{\mu} = 0, \; R = \sigma^2 I$: iid zero mean

$|R| = (\sigma^2)^n$

$$\left.\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \cdot e^{-\|\bar{x}\|^2/2\sigma^2}\right\}$$ iid zero mean.

Gaussian Classification

- Recall, the PDF of a ***Gaussian random vector*** is given as

$$f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) = \frac{1}{\sqrt{(2\pi)^n|\mathbf{R}|}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})^T\mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}})}$$

Multivariate Gaussian PDF.

Cramer-Rao Bound

- The ***mean*** and ***covariance matrix*** are defined as

$$E\{\bar{x}\} = \bar{\mu}$$

$$E\{(\bar{x}-\bar{\mu})(\bar{x}-\bar{\mu})^T\} = R.$$

Gaussian Classification

- The ***mean*** and ***covariance matrix*** are defined as

$$E\{\bar{\mathbf{x}}\} = \bar{\boldsymbol{\mu}}$$

$$E\{(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}})(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}})^T\} = \mathbf{R}$$

Gaussian Classification

2 Gaussian classes.

- Classification with **Gaussian classes**.

$C_0 : \bar{\mu}_0, R$

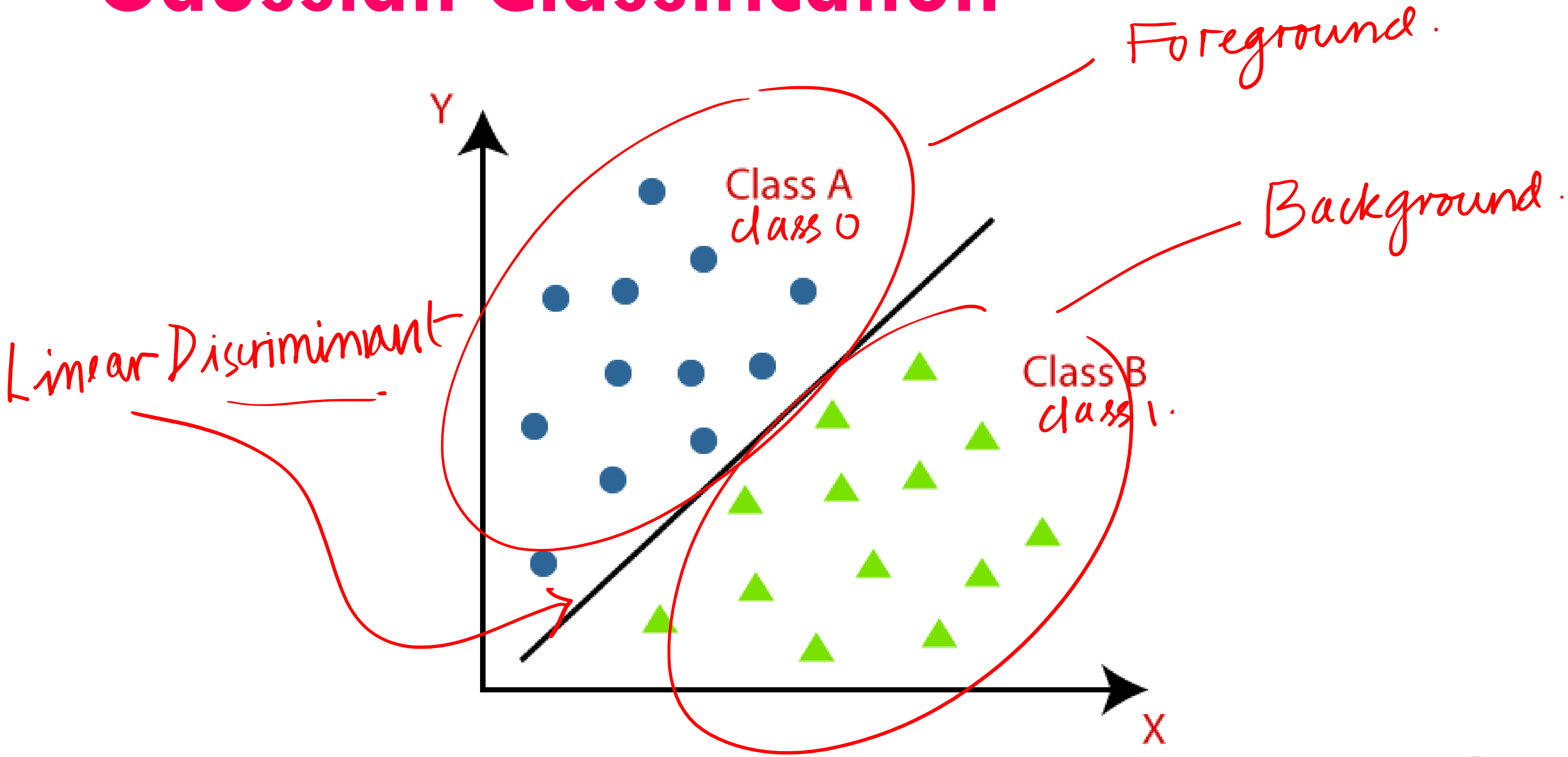
$C_1 : \bar{\mu}_1, R$

C_0.

C_1.

- Example: **Foreground pixels**, **Background pixels**.

Gaussian Classification



Gaussian Classification

- Consider the input vectors $\bar{\mathbf{x}}$ drawn from ***two Gaussian classes***

$$\left.\begin{aligned} \mathcal{C}_0 &: \mathcal{N}(\bar{\mu}_0, R). \\ \mathcal{C}_1 &: \mathcal{N}(\bar{\mu}_1, R). \end{aligned}\right\}$$

ML: Gaussian Discriminant Analysis.

Gaussian Classification

- Consider the input vectors $\bar{\mathbf{x}}$ drawn from ***two Gaussian classes***

$$\mathcal{C}_0\text{: Mean } \bar{\boldsymbol{\mu}}_0 \text{ and covariance } \mathbf{R}$$
$$\mathcal{C}_1\text{: Mean } \bar{\boldsymbol{\mu}}_1 \text{ and covariance } \mathbf{R}$$

- Also termed **Gaussian Discriminant Analysis**

$|R| =$ determinant of R.

Gaussian Classification

Likelihood of class 0.

- Thus, the ***likelihoods*** of the two classes are

$$p(\bar{\mathbf{x}}; \mathcal{C}_0) = \frac{1}{\sqrt{(2\pi)^n |R|}} e^{-\frac{1}{2}(\bar{x}-\bar{\mu}_0)^T R^{-1} (\bar{x}-\bar{\mu}_0)}$$

$$p(\bar{\mathbf{x}}; \mathcal{C}_1) = \frac{1}{\sqrt{(2\pi)^n |R|}} \cdot e^{-\frac{1}{2}(\bar{x}-\bar{\mu}_1)^T R^{-1} (\bar{x}-\bar{\mu}_1)}$$

likelihood of class 1.

Gaussian Classification

- Thus, the ***likelihoods*** of the two classes are

$$p(\bar{\mathbf{x}};\mathcal{C}_0)=\frac{1}{\sqrt{(2\pi)^n|\mathbf{R}|}}e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_0)^T\mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_0)}$$

$$p(\bar{\mathbf{x}};\mathcal{C}_1)=\frac{1}{\sqrt{(2\pi)^n|\mathbf{R}|}}e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_1)^T\mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_1)}$$

Gaussian Classification

- Choose the class that ***maximizes the likelihood*** (ML).
- Therefore, **choose** $\mathcal{C}_0$ if

choose class that has highest likelihood.

$$p(\bar{x}; C_0) \geq p(\bar{x}; C_1).$$

choose C_1 if $p(\bar{x}; C_0) < p(\bar{x}; C_1)$

Gaussian Classification

- Choose the class that ***maximizes the likelihood***
- Therefore, **choose** $\mathcal{C}_0$ if

$$p(\bar{\mathbf{x}}; \mathcal{C}_0) \geq p(\bar{\mathbf{x}}; \mathcal{C}_1)$$

Gaussian Classification

- This can be simplified as

H_0:

choose C_0 if

$$\cancel{\frac{1}{\sqrt{(2\pi)^n|R|}}} e^{-\frac{1}{2}(\bar{x}-\bar{\mu}_0)^T R^{-1}(\bar{x}-\bar{\mu}_0)} \geq \cancel{\frac{1}{\sqrt{(2\pi)^n|R|}}} e^{-\frac{1}{2}(\bar{x}-\bar{\mu}_1)^T R^{-1}(\bar{x}-\bar{\mu}_1)}.$$

$$\Rightarrow (\bar{x}-\bar{\mu}_0)^T R^{-1}(\bar{x}-\bar{\mu}_0) \leq (\bar{x}-\bar{\mu}_1)^T R^{-1}(\bar{x}-\bar{\mu}_1)$$

Gaussian Classification

$$\Rightarrow \cancel{\bar{x}^T R^{-1} \bar{x}} - 2\bar{\mu}_0^T R^{-1} \bar{x} + \bar{\mu}_0^T R^{-1} \bar{\mu}_0$$

$$< \cancel{\bar{x}^T R^{-1} \bar{x}} - 2\bar{\mu}_1^T R^{-1} \bar{x} + \bar{\mu}_1^T R^{-1} \bar{\mu}_1$$

$$\Rightarrow (\bar{\mu}_0 - \bar{\mu}_1)^T R^{-1} \bar{x} + \frac{1}{2}\left(\bar{\mu}_1^T R^{-1} \bar{\mu}_1 - \bar{\mu}_0^T R^{-1} \bar{\mu}_0\right) \geq 0$$

$$\Rightarrow (\bar{\mu}_0 - \bar{\mu}_1)^T R^{-1} \left(\bar{x} - \left(\frac{\bar{\mu}_0 + \bar{\mu}_1}{2}\right)\right) \geq 0$$

Gaussian Classification

choose C_0 if

$$\bar{h}^T(\bar{x} - \hat{\mu}) \geq 0$$

N dimensional plane.

$$\bar{h} = R^{-1}(\bar{\mu}_0 - \bar{\mu}_1)$$

$$\tilde{\mu} = \frac{\bar{\mu}_0 + \bar{\mu}_1}{2}$$

choose C_1 if $\bar{h}^T(\bar{x} - \hat{\mu}) < 0$.

Gaussian Classification

- This can be simplified as

$$p(\bar{\mathbf{x}}; \mathcal{C}_0) \geq p(\bar{\mathbf{x}}; \mathcal{C}_1)$$

$$\Rightarrow \frac{1}{\sqrt{(2\pi)^n |\mathbf{R}|}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_0)}$$

$$\geq \frac{1}{\sqrt{(2\pi)^n |\mathbf{R}|}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_1)}$$

$$\Rightarrow (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_0) \leq (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1} (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_1)$$

Gaussian Classification

- Further

$$(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_0) \leq (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1} (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_1)$$

$$\Rightarrow \bar{\mathbf{x}} \mathbf{R}^{-1} \bar{\mathbf{x}} - 2 \bar{\boldsymbol{\mu}}_0^T \mathbf{R}^{-1} \bar{\mathbf{x}} + \bar{\boldsymbol{\mu}}_0^T \mathbf{R}^{-1} \bar{\boldsymbol{\mu}}_0 \leq$$

$$\bar{\mathbf{x}} \mathbf{R}^{-1} \bar{\mathbf{x}} - 2 \bar{\boldsymbol{\mu}}_1^T \mathbf{R}^{-1} \bar{\mathbf{x}} + \bar{\boldsymbol{\mu}}_1^T \mathbf{R}^{-1} \bar{\boldsymbol{\mu}}_1$$

$$\Rightarrow (\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1} \bar{\boldsymbol{x}}$$

$$+ \frac{\mathbf{1}}{\mathbf{2}} (\bar{\boldsymbol{\mu}}_1^T \mathbf{R}^{-1} \bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0^T \mathbf{R}^{-1} \bar{\boldsymbol{\mu}}_0) \geq \mathbf{0}$$

Gaussian Classification

- Further

$$\Rightarrow (\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1} \left(\bar{\mathbf{x}} - \frac{1}{2}(\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1) \right) \geq \mathbf{0}$$

$$\Rightarrow \bar{\mathbf{h}}^T (\bar{\mathbf{x}} - \tilde{\boldsymbol{\mu}}) \geq 0$$

$$\underbrace{\tilde{\boldsymbol{\mu}} = \frac{1}{2}(\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1)}_{\text{midpt of } \bar{\mu}_0, \bar{\mu}_1}, \quad \bar{\mathbf{h}} = \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)$$

Gaussian Classification

Another way to write same classifier

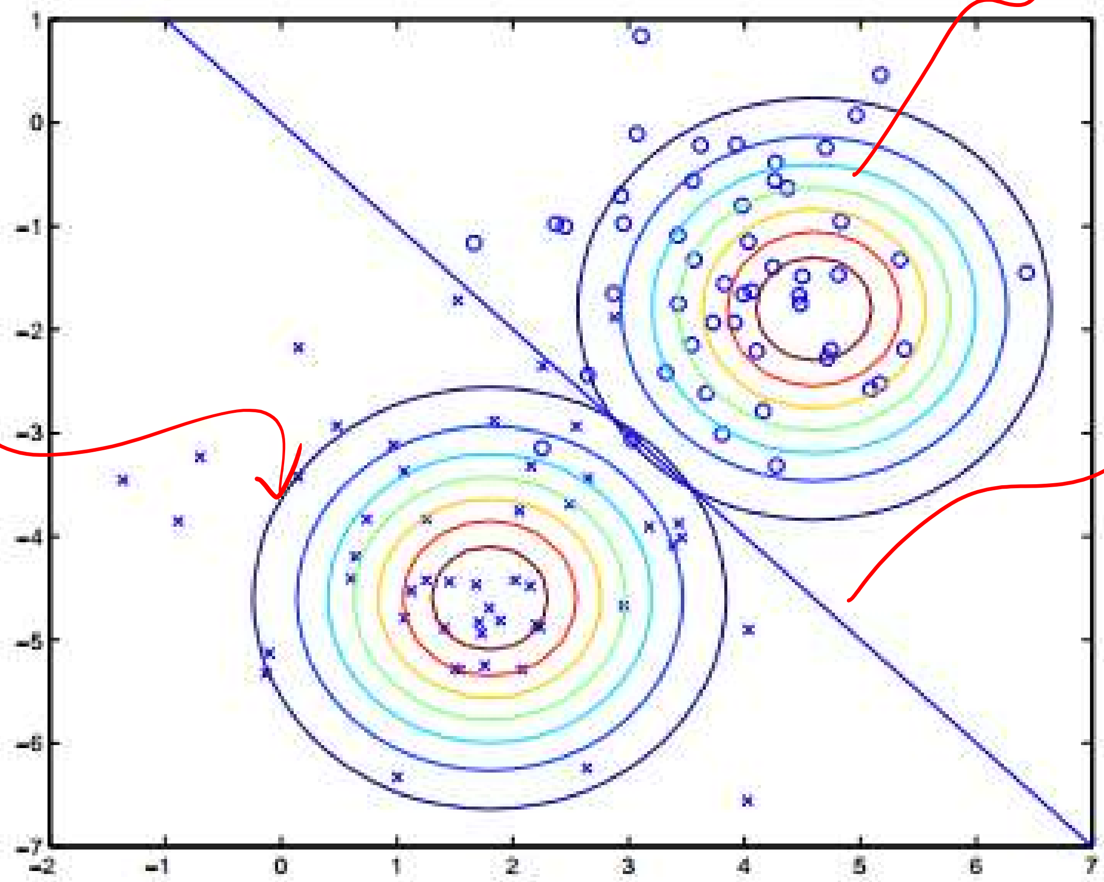
- Also

choose C_0 if $-(\bar{\mu}_0-\bar{\mu}_1)^T R^{-1}(\bar{x}-\tilde{\mu}) \lessgtr 0$

$$\Rightarrow (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}\left(\bar{\mathbf{x}} - \frac{1}{2}(\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1)\right) \leq \mathbf{0}$$

$$\Rightarrow (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}\bar{\mathbf{x}} \leq \frac{1}{2}(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0)$$

Gaussian Classification



Class 0 ·

Class 1

Linear Discriminant

Separates classes.

Special Case

- Consider the ***special case*** $\mathbf{R} = \sigma^2 \mathbf{I}$
- It follows that

$$\bar{\mathbf{h}} = R^{-1}(\bar{\mu}_0 - \mu_1) = \frac{1}{\sigma^2}(\bar{\mu}_0 - \bar{\mu}_1).$$

Special Case

- Consider the *special case* $\mathbf{R} = \sigma^2\mathbf{I}$
- It follows that

$$\bar{\mathbf{h}} = \frac{1}{\sigma^2}\mathbf{I}(\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1) \sim (\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)$$

Special Case

- The **hyperplane** reduces to

$$\frac{\bar{\mu}_0 + \bar{\mu}_1}{2}$$

C_0 if

$$\frac{1}{\sigma^2}(\bar{\mu}_0 - \bar{\mu}_1)^T(\bar{x} - \tilde{\mu}) \geq 0$$

$$\Rightarrow (\bar{\mu}_0 - \bar{\mu}_1)^T(\bar{x} - \hat{\mu}) \geq 0.$$

$$\Rightarrow \underline{(\bar{\mu}_0 - \bar{\mu}_1)^T\left(\bar{x} - \left(\frac{\bar{\mu}_0 + \bar{\mu}_1}{2}\right)\right) \geq 0.}$$

Special Case

- The **hyperplane** reduces to

$$(\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)^T \left(\bar{\mathbf{x}} - \frac{1}{2}(\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1) \right) \geq \mathbf{0}$$

Special Case

- Pictorially

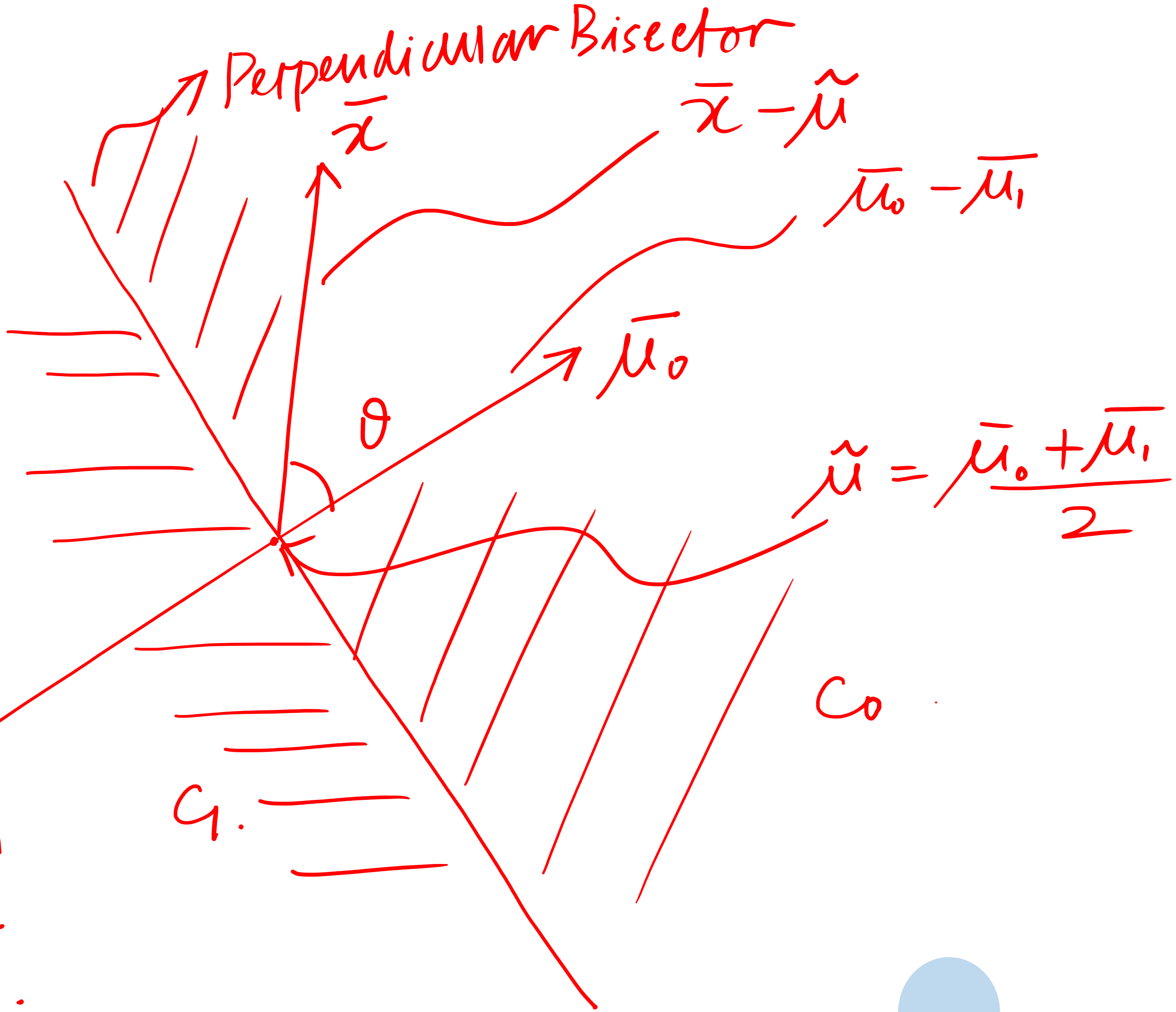
$$(\bar{\mu}_0-\bar{\mu}_1)^T(\bar{x}-\tilde{\mu})$$

$$\Rightarrow \|\bar{\mu}_0-\bar{\mu}_1\|\,\|\bar{x}-\tilde{\mu}\|.\cos\theta \geq 0$$

$$\Rightarrow -90^\circ \leq \theta \leq 90^\circ$$

$$\left.\begin{array}{l} 180^\circ \geq \theta > 90^\circ \\ -180 \leq \theta < -90^\circ \end{array}\right\} C_1 :$$

Nearest Neighbor decoding rule.



Special Case

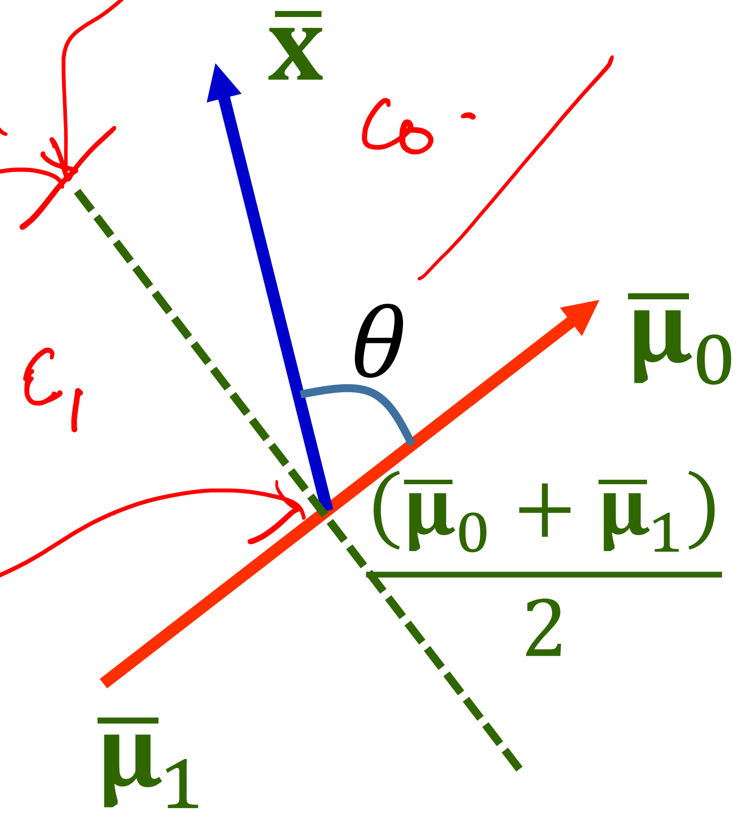
- Pictorially

Linear discriminant

Nearest ~~Neighbor~~ decoding rule.

perpendicular bisector of $\bar{\mu}_0, \bar{\mu}_1$

midpoint.



$-90^\circ \leq \theta \leq 90^\circ \Rightarrow C_0$.

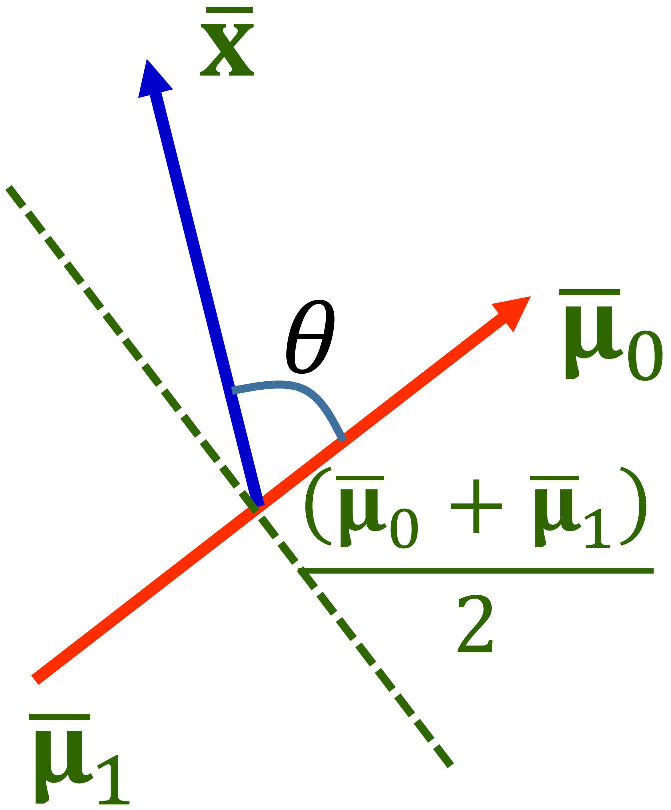
Dot product
$(\bar{\mu}_0 - \bar{\mu}_1)^T(\bar{x} - \tilde{\mu}) \geq 0$
as $\cos\theta \geq 0$.

$90^\circ < \theta < 270^\circ \Rightarrow C_1$
$\cos\theta < 0$
$\Rightarrow$ Dot product
$(\bar{\mu}_0 - \bar{\mu}_1)^T(\bar{x} - \tilde{\mu}) < 0$.

Special Case

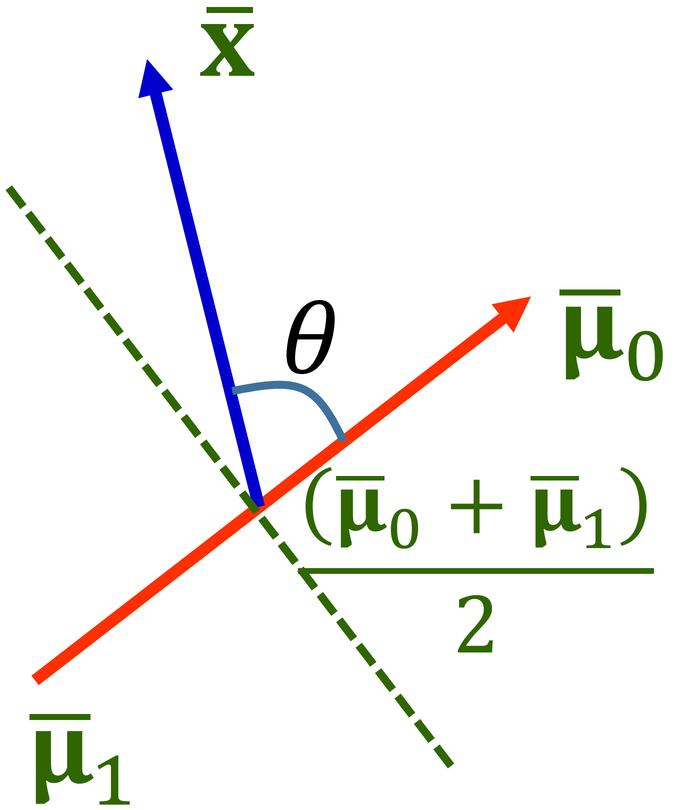
- It can be seen that $(\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)^T \left(\bar{\mathbf{x}} - \frac{1}{2}(\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1) \right) \geq 0$,
- when $-90^0 \leq \theta \leq 90^0$ since $\cos\theta \geqslant 0$.

$\Rightarrow$ inner product $\geqslant 0$.



Special Case

- Thus, the hyperplane is the ***perpendicular bisector*** of $\bar{\boldsymbol{\mu}}_0$ and $\bar{\boldsymbol{\mu}}_1$



Example

- Determine the classifier for the **Gaussian classification** problem
- with the two classes C_0, C_1 distributed as

$$C_0 \sim N\left(\begin{bmatrix} 2 \\ -4 \end{bmatrix}, \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{8} \end{bmatrix}\right), C_1 \sim N\left(\begin{bmatrix} -4 \\ 2 \end{bmatrix}, \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{8} \end{bmatrix}\right)$$

$\bar{\mu}_0$ $\bar{\mu}_1$ R

$$R^{-1} = \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}$$

Example

- Classifier chooses $\mathcal{H}_0$ if

$$(\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1} \left(\bar{\mathbf{x}} - \frac{\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1}{2} \right) \geq 0$$

$\Rightarrow$

$$\bar{\mu}_0 - \bar{\mu}_1 = \begin{bmatrix} 2 \\ -4 \end{bmatrix} - \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

$$\frac{\bar{\mu}_0 + \bar{\mu}_1}{2} = \frac{1}{2}\left(\begin{bmatrix} 2 \\ -4 \end{bmatrix} + \begin{bmatrix} -4 \\ 2 \end{bmatrix} \right)$$

$$= \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

Example

- Classifier chooses $\mathcal{H}_0$ if $x_1 - 2x_2 \geq 1$

$$\begin{bmatrix} 6 & -6 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix} \left(\bar{x} - \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right) \geq 0$$

$$\Rightarrow \begin{bmatrix} 6 & -6 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) \geq 0$$

$$\Rightarrow 24x_1 - 48x_2 + 24 - 48 \geq 0$$

$$\Rightarrow 24x_1 - 48x_2 \geq 24$$

$$\Rightarrow \underline{x_1 - 2x_2 \geq 1}$$

Example

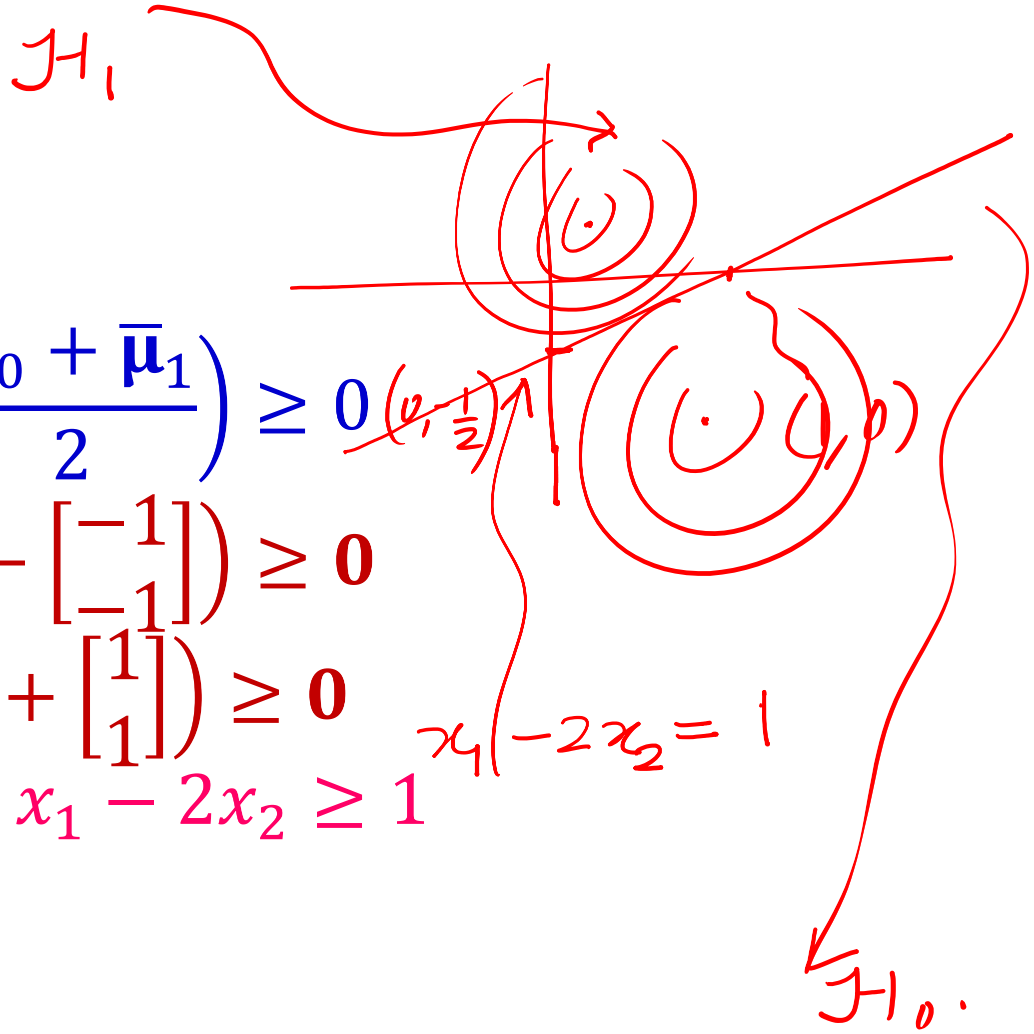
- Classifier chooses $\mathcal{H}_0$ if

$$(\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)^T \boldsymbol{\Sigma}^{-1}\left(\bar{\mathbf{x}} - \frac{\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1}{2}\right) \geq 0$$

$$\Rightarrow \left(\begin{bmatrix} 6 \\ -6 \end{bmatrix}\right)^T \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}\left(\bar{\mathbf{x}} - \begin{bmatrix} -1 \\ -1 \end{bmatrix}\right) \geq \mathbf{0}$$

$$\Rightarrow [6 \quad -6] \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}\left(\bar{\mathbf{x}} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \geq \mathbf{0}$$

$$\Rightarrow 24x_1 - 48x_2 \geq 24 \Rightarrow x_1 - 2x_2 \geq 1$$



Analysis

- Classifier chooses $\mathcal{H}_0$ if

Linear Discriminant Function.

C_0/H_0 if $x_1 - 2x_2 \geq 1$

C_1/H_1 if $x_1 - 2x_2 < 1$

Analysis

P_D, P_{FA} ?

- Classifier chooses $\mathcal{H}_0$ if

$$(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}} \leq \frac{\mathbf{1}}{\mathbf{2}} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0)$$

Analysis

- Classifier chooses $\mathcal{H}_1$ if

Analysis

- Classifier chooses $\mathcal{H}_1$ if

$$(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}} > \frac{1}{2} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0)$$

Analysis

- What is P_{FA} Probability of False Alarm under $\mathcal{H}_0$, what is probability decision $= \mathcal{H}_1$?
- Under $\mathcal{H}_0$, $\bar{\mathbf{x}} \sim \mathcal{N}(\bar{\boldsymbol{\mu}}_0, \mathbf{R})$

Gaussian RV since this is a Linear transformation of Gaussian

$$\Rightarrow (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}} \sim E\{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} \bar{x}\}$$

$$= (\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} E\{\bar{x}\}$$

$$= (\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} \bar{\mu}_0$$

Expected value under $\mathcal{H}_0$.

Analysis

$$E\left\{(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{x}-\bar{\mu}_0)(\bar{x}-\bar{\mu}_0)^T R^{-1}(\bar{\mu}_1-\bar{\mu}_0)\right\}$$

$$= (\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}\underbrace{E\left\{(\bar{x}-\bar{\mu}_0)(\bar{x}-\bar{\mu}_0)^T\right\}}_{R} R^{-1}(\bar{\mu}_1-\bar{\mu}_0)$$

$$= (\bar{\mu}_1-\bar{\mu}_0)^T \cancel{R^{-1}}\,\cancel{R}\cdot R^{-1}(\bar{\mu}_1-\bar{\mu}_0)$$

$$= (\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{\mu}_1-\bar{\mu}_0).$$

covariance of $(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}\bar{x}$

Analysis

- What is P_{FA}
- Under $\mathcal{H}_0$, $\bar{\mathrm{x}} \sim \mathcal{N}(\bar{\boldsymbol{\mu}}_0, \mathbf{R})$

Gaussian Random vector under $\mathcal{H}_0$.

$$\Rightarrow (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}}$$

$$\sim \mathcal{N}\left((\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\boldsymbol{\mu}}_0, (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)\right)$$

Mean

Variance.

Analysis

$$\Pr(X > \alpha) = \Pr\left(\frac{X-\mu}{\sigma} > \frac{\alpha-\mu}{\sigma}\right) = Q\left(\frac{\alpha-\mu}{\sigma}\right)$$

standard Gaussian RV

- FA occurs when Under $\mathcal{H}_0$ if decision is $\mathcal{H}_1$

$$P_{FA} = \Pr\left((\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} \bar{x} > \frac{1}{2}(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1}(\bar{\mu}_0 + \bar{\mu}_1)\right)$$

$$= Q\left(\frac{\frac{1}{2}(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1}(\bar{\mu}_0 + \bar{\mu}_1) - (\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1}\bar{\mu}_0}{\sqrt{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1}(\bar{\mu}_1 - \bar{\mu}_0)}}\right)$$

Special case

- FA occurs when Under $\mathcal{H}_0$ if decision is $\mathcal{H}_1$

$$(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}} > \frac{1}{2} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0)$$

Special case

- P_{FA} is given as

$$\Pr\left((\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}} > \frac{1}{2} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0) \right)$$

$$= Q\left(\frac{\frac{1}{2}(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 + \bar{\mu}_0) - (\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} \mu_0}{\sqrt{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 - \bar{\mu}_0)}} \right)$$

$$= Q\left(\frac{\frac{1}{2}(\bar{\mu}_1 - \bar{\mu}_0) R^{-1} (\bar{\mu}_1 - \bar{\mu}_0)}{\sqrt{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 - \bar{\mu}_0)}} \right)$$

$$= Q\left(\frac{1}{2} \sqrt{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 - \bar{\mu}_0)} \right).$$

Special case

$R = \sigma^2 I$

P_{FA} = Probability of False Alarm for Gaussian classification

- P_{FA} is given as

$$\Pr\left((\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}} > \frac{1}{2} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0) \right)$$

$$= Q\left(\frac{\frac{1}{2} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0) - (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\boldsymbol{\mu}}_0}{\sqrt{(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)}} \right)$$

$$= Q\left(\frac{1}{2} \sqrt{(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)} \right)$$

$$= Q\left(\frac{1}{2} \sqrt{(\bar{\mu}_1 - \bar{\mu}_0)^T \frac{1}{\sigma^2} I (\bar{\mu}_1 - \bar{\mu}_0)} \right)$$

$$= Q\left(\frac{\|\bar{\mu}_1 - \bar{\mu}_0\|}{2\sigma} \right)$$

Analysis

under $\mathcal{H}_1$, what is probability that decision $= \mathcal{H}_1$?

- What is P_D
- Under $\mathcal{H}_1$, $\bar{\mathbf{x}} \sim \mathcal{N}(\bar{\boldsymbol{\mu}}_1, \mathbf{R})$

Gaussian Because linear transformation of Gaussian

SCALAR QUANTITY.

$$\Rightarrow (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}} \sim E\{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} \bar{x}\}$$

$$= (\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} E\{\bar{x}\}$$

$$= (\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} \bar{\mu}_1$$

Mean or Expected value.

Analysis

$$E\left\{(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{x}-\bar{\mu}_1)(\bar{x}-\bar{\mu}_1)^T R^{-1}(\bar{\mu}_1-\bar{\mu}_0)\right\}$$

$$=(\bar{\mu}_1-\bar{\mu}_0)^T \cancel{R^{-1}}\,\cancel{R}\,R^{-1}(\bar{\mu}_1-\bar{\mu}_0)\}$$

$$=(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{\mu}_1-\bar{\mu}_0)$$

variance.

Analysis

- What is P_D
- Under $\mathcal{H}_1$, $\bar{\mathrm{x}} \sim \mathcal{N}(\bar{\mu}_1, \mathbf{R})$

Mean

Variance

$$\Rightarrow (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}}$$

$$\sim \mathcal{N}\left((\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\boldsymbol{\mu}}_1, (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)\right)$$

Analysis

- Detection occurs when Under $\mathcal{H}_1$

$$(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} \bar{x} > \frac{1}{2} (\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 + \bar{\mu}_0).$$

Special case

- Detection occurs when Under $\mathcal{H}_1$

$$(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} \bar{\mathbf{x}} > \frac{\mathbf{1}}{\mathbf{2}} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0)$$

Special case

- P_D is given as

$$\Pr\left((\bar{\boldsymbol{\mu}}_1-\bar{\boldsymbol{\mu}}_0)^T\mathbf{R}^{-1}\bar{\mathbf{x}} > \overbrace{\frac{1}{2}(\bar{\boldsymbol{\mu}}_1-\bar{\boldsymbol{\mu}}_0)^T\mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1+\bar{\boldsymbol{\mu}}_0)}^{\alpha}\right)$$

$$= Q\left(\frac{\frac{1}{2}(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{\mu}_1+\bar{\mu}_0) - \underbrace{(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}\mu_1}_{\text{mean}}}{\sqrt{(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\mu_1-\bar{\mu}_0)}}\right)$$

$$= Q\left(\frac{-\frac{1}{2}(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{\mu}_1-\bar{\mu}_0)}{\sqrt{(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{\mu}_1-\bar{\mu}_0)}}\right)$$

$$= \underbrace{Q\left(-\frac{1}{2}\sqrt{(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{\mu}_1-\bar{\mu}_0)}\right)}_{P_D}$$

Special case

P_D = Probability of Detection for Gaussian classification

- P_D is given as

$$\Pr\left((\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}\bar{\mathbf{x}} > \frac{1}{2}(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0)\right)$$

$$= Q\left(\frac{\frac{1}{2}(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1 + \bar{\boldsymbol{\mu}}_0) - (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}\bar{\boldsymbol{\mu}}_1}{\sqrt{(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)}}\right)$$

$$= Q\left(-\frac{1}{2}\sqrt{(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)}\right)$$

Special case

- P_{MD} is given as

$$1 - Q(-x) = Q(x).$$

$$P_{MD} = 1 - P_D = 1 - Q\left(-\frac{1}{2}\sqrt{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 - \bar{\mu}_0)}\right)$$

$$= Q\left(\frac{1}{2}\sqrt{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 - \bar{\mu}_0)}\right) = P_{FA}$$

Special case

- P_{MD} is given as

$$= Q\left(\frac{1}{2}\sqrt{(\overline{\boldsymbol{\mu}}_1 - \overline{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\overline{\boldsymbol{\mu}}_1 - \overline{\boldsymbol{\mu}}_0)}\right) = P_{FA}.$$

Special case

- P_e is given as

$$\frac{1}{2}P_{FA} + \frac{1}{2}P_{MD} = Q\left(\frac{1}{2}\sqrt{(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)}\right)$$

Probability of Error for ~~Gaussian~~ classification

Mean of class 1

R = covariance matrix

Mean of class C_0

Example

- For the **Gaussian classification** problem with classes, what is P_e

$$C_0 \sim N\left(\begin{bmatrix} 2 \\ -4 \end{bmatrix}, \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{8} \end{bmatrix}\right), C_1 \sim N\left(\begin{bmatrix} -4 \\ 2 \end{bmatrix}, \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{8} \end{bmatrix}\right)$$

$\bar{\mu}_0$ R $\bar{\mu}_1$ R

What is P_e?

$$\bar{\mu}_1 - \bar{\mu}_0 = \begin{bmatrix} -6 \\ 6 \end{bmatrix} \quad R^{-1} = \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}$$

Example

- The P_e is

Probability of Error

$$Q\left(\frac{1}{2}\sqrt{(\bar{\mu}_1-\bar{\mu}_0)^T R^{-1}(\bar{\mu}_1-\bar{\mu}_0)}\right)$$

$$= Q\left(\frac{1}{2}\sqrt{\begin{bmatrix}-6 & 6\end{bmatrix}\begin{bmatrix}4 & 0\\ 0 & 8\end{bmatrix}\begin{bmatrix}-6\\ 6\end{bmatrix}}\right)$$

$$= Q\left(\frac{1}{2}\sqrt{36\times4+36\times8}\right)$$

$$= Q\left(\frac{1}{2}\sqrt{36\times12}\right) = Q\left(\cancel{\frac{1}{2}}\times6\times\cancel{2}\sqrt{3}\right)$$

$$= Q(6\sqrt{3})$$

Example

Probability of Error for Gaussian classification problem

- The P_e is

$$Q\left(\frac{1}{2}\sqrt{(\bar{\boldsymbol{\mu}}_1-\bar{\boldsymbol{\mu}}_0)^T\mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1-\bar{\boldsymbol{\mu}}_0)}\right)$$

$$= Q\left(\frac{1}{2}\sqrt{\begin{bmatrix}\mathbf{-6} & \mathbf{6}\end{bmatrix}\begin{bmatrix}4 & 0\\ 0 & 8\end{bmatrix}\begin{bmatrix}\mathbf{-6}\\ 6\end{bmatrix}}\right)$$

$$= Q\left(\frac{1}{2}\sqrt{36\times 4+36\times 8}\right) = Q\left(\frac{1}{2}\sqrt{36\times 12}\right)$$

$$= Q\left(6\sqrt{3}\right)$$

Optimal Signaling

- How to find **optimal signals** $\bar{\boldsymbol{\mu}}_1, \bar{\boldsymbol{\mu}}_0$?
- Note P_e is given as

$$P_e = Q\left(\frac{1}{2}\sqrt{(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 - \bar{\mu}_0)}\right)$$

Optimal Signaling

Signal Design problem.

- How to find optimal signals $\bar{\boldsymbol{\mu}}_1, \bar{\boldsymbol{\mu}}_0$?

How to choose?

minimize Pe

- Note P_e is given as

$$Q\left(\frac{1}{2}\sqrt{(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1} (\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)}\right)$$

maximize.

Decreasing Function

$\Rightarrow$ To minimize Pe

maximize $(\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 - \bar{\mu}_0)$.

Optimal Signaling

- To minimize P_e

$$\max \; (\bar{\mu}_1 - \bar{\mu}_0)^T R^{-1} (\bar{\mu}_1 - \bar{\mu}_0).$$

- Let $\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0 = \bar{\mathbf{s}}$. Definition

$$\max \; \bar{s}^T R^{-1} \bar{s}$$

Optimal Signaling

- To minimize P_e

$$\max(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0)$$

- Let $\bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0 = \bar{\mathbf{s}}$.

$$\max \bar{\mathbf{s}}^T \mathbf{R}^{-1} \bar{\mathbf{s}}$$

Optimal Signaling

application of eigenvalue decomposition.

- Let $\mathbf{R}$ have the **eigenvalue decomposition**

$$R = U \Lambda U^T$$

Diagonal matrix of eigenvalues.

eigenvectors.

$$= \underbrace{\begin{bmatrix} \bar{u}_1 & \bar{u}_2 & \cdots & \bar{u}_N \end{bmatrix} \overbrace{\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{bmatrix}}^{\Lambda} \begin{bmatrix} \bar{u}_1^T \\ \bar{u}_2^T \\ \vdots \\ \bar{u}_N^T \end{bmatrix}}_{\text{Eigenvalue Decomposition}}$$

Optimal Signaling

- Let $\mathbf{R}$ have the **eigenvalue decomposition**

Property of Eigenvalue/Eigenvector

$$\mathbf{R}\bar{u}_i = \lambda_i \bar{u}_i$$

Eigenvalue.

Eigenvector

$$|\mathbf{R} - \lambda I| = 0$$

Determinant.

Optimal Signaling

- Let $\mathbf{R}$ have the **eigenvalue decomposition**

Diagonal matrix of eigenvalues.

$$\mathbf{R} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$$

$$= [\bar{\mathbf{u}}_1 \quad \cdots \quad \bar{\mathbf{u}}_N] \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{bmatrix} \begin{bmatrix} \bar{\mathbf{u}}_1^T \\ \vdots \\ \bar{\mathbf{u}}_N^T \end{bmatrix}$$

Optimal Signaling

- Note the following properties

covariance matrix

$$R = U \Lambda U^T$$

Positive Definite matrix

Eigenvalues > 0.

$\bar{u}_i^T \bar{u}_j = 0$ if $i \neq j$

$\|\bar{u}_i\|^2 = 1$.

Therefore: $\lambda_i > 0$

Eigenvectors orthonormal.

$$U U^T = U^T U = I$$

U is a unitary matrix

Optimal Signaling

- Note the following properties

U = unitary

$$\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$$

Eigenvalues > 0

$$\lambda_i > 0$$

R is PD matrix.

Optimal Signaling

- Let $\bar{s}$ be

Expanding $\bar{s}$ using the basis of orthonormal eigenvectors of R.

$$\bar{\mathbf{s}} = \mathbf{U}\bar{\boldsymbol{\alpha}}$$

$$= \begin{bmatrix} \bar{u}_1 & \bar{u}_2 & \cdots & \bar{u}_N \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{bmatrix} = \alpha_1 \bar{u}_1 + \cdots + \alpha_N \bar{u}_N$$

$$= \sum_i \alpha_i \bar{u}_i .$$

Optimal Signaling

- Let $\bar{\mathbf{s}}$ be

$$\bar{\mathbf{s}} = \mathbf{U}\bar{\boldsymbol{\alpha}}$$
$$= \alpha_1 \bar{\mathbf{u}}_1 + \cdots + \alpha_N \bar{\mathbf{u}}_N = \sum_i \alpha_i \bar{u}_i$$

Optimal Signaling

- Furthermore, $\mathbf{R}^{-1}$ is given as

$$\mathbf{R}^{-1} = U\Lambda^{-1}U^{T} = U\begin{bmatrix} \frac{1}{\lambda_1} & & & \\ & \frac{1}{\lambda_2} & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_N} \end{bmatrix}U^{T}$$

Λ^{-1} is diagonal matrix of eigenvalues of R^{-1}

$$R^{-1} \cdot R = U\Lambda^{-1}\cancel{U^{T}}\cancel{U}\Lambda U^{T}$$

$$= U\cancel{\Lambda^{-1}}\cancel{\Lambda}U^{T}$$

$$= UU^{T} = I$$

Optimal Signaling

- Furthermore, $\mathbf{R}^{-1}$ is given as

Eigenvectors of R, R⁻¹ are same.

$$\mathbf{R}^{-1} = \mathbf{U}\boldsymbol{\Lambda}^{-1}\mathbf{U}^T$$

$$= \begin{bmatrix} \bar{\mathbf{u}}_1 & \cdots & \bar{\mathbf{u}}_N \end{bmatrix} \begin{bmatrix} \lambda_1^{-1} & & \\ & \ddots & \\ & & \lambda_N^{-1} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{u}}_1^T \\ \vdots \\ \bar{\mathbf{u}}_N^T \end{bmatrix}$$

Optimal Signaling

- One can formulate the **optimization problem** as

constrained optimization problem.

$$\max: \quad \bar{s}^T R^{-1} \bar{s}$$

$$\text{s.t.} \quad \|\bar{s}\|^2 = 1$$

constraint

increase in unbounded fashion if $\|\bar{s}\|$ increases.

Objective function

Optimal Signaling

- One can formulate the **optimization problem** as

$$\max \bar{\mathbf{s}}^T \mathbf{R}^{-1} \bar{\mathbf{s}}$$
$$\|\bar{\mathbf{s}}\|^2 = 1$$

Optimal Signaling

$\bar{s} = U\bar{\alpha} \Rightarrow \bar{s}^T = \bar{\alpha}^T U^T$

- Simplification

$$\bar{\mathbf{s}}^T \mathbf{R}^{-1} \bar{\mathbf{s}} = \bar{\alpha}^T \cancel{U^T} \cdot \cancel{U} \Lambda^{-1} \cancel{U^T} \cancel{U} \bar{\alpha}$$

$$= \bar{\alpha}^T \cdot \Lambda^{-1} \bar{\alpha} = [\alpha_1 \; \cdots \; \alpha_N] \begin{bmatrix} 1/\lambda_1 & & \\ & \ddots & \\ & & 1/\lambda_N \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}$$

$$= \sum_i \frac{\alpha_i^2}{\lambda_i}$$

Optimal Signaling

- Simplification

$$\bar{\mathbf{s}}^T \mathbf{R}^{-1} \bar{\mathbf{s}} = \bar{\boldsymbol{\alpha}}^T \mathbf{U}^T \mathbf{U} \boldsymbol{\Lambda}^{-1} \mathbf{U}^T \mathbf{U} \bar{\boldsymbol{\alpha}}$$

$$= \bar{\boldsymbol{\alpha}}^T \boldsymbol{\Lambda}^{-1} \bar{\boldsymbol{\alpha}} = \sum_i \frac{\alpha_i^2}{\lambda_i}$$

Optimal Signaling

- Simplification

$$\|\bar{\mathbf{s}}\|^2 = \quad \bar{s}^T \bar{s} = \bar{\alpha}^T \cancel{U^T} \cancel{U} \bar{\alpha}$$

$$= \bar{\alpha}^T \cdot \bar{\alpha} = \sum_i \alpha_i^2 = \|\bar{\alpha}\|^2$$

Optimal Signaling

- Simplification

$$\|\bar{\mathbf{s}}\|^2 = \bar{\boldsymbol{\alpha}}^T \mathbf{U}^T \mathbf{U} \bar{\boldsymbol{\alpha}} = \sum_i \alpha_i^2$$

Optimal Signaling

- **Optimization problem** can be modified as

$$\max: \quad \bar{s}^T R^{-1} \bar{s} \equiv \sum_i \frac{\alpha_i^2}{\lambda_i}$$

$$\text{s.t.} \quad \|\bar{s}\|^2 = 1 \equiv \sum_i \alpha_i^2 = 1$$

Optimal Signaling

- Optimization problem can be modified as

$$\max \sum_i \frac{\alpha_i^2}{\lambda_i}$$

$$\sum_i \alpha_i^2 = 1$$

Optimal Signaling

- Solution is

$$\max \sum_i \alpha_i^2 / \lambda_i$$

$$\text{st. } \sum_i \alpha_i^2 = 1$$

set $\alpha_i^2 = 1$ for i such that λ_i is min!

Rest $\alpha_i^2 = 0$.

Optimal Signaling

$\bar{u}_i =$ orthonormal eigenvector of R, corresponding to min eigenvalue.

- Solution is

$$\alpha_i = 1, i = \arg\min_j \lambda_j \Rightarrow \bar{\mathbf{s}} = \bar{\mathbf{u}}_i$$

$$\alpha_i = 0, \text{otherwise}$$

- $\bar{\mathbf{u}}_i$ is eigenvector corresponding to min eigenvalue

Optimal Signaling

R = noise covariance

$\bar{\mu}_0 + \bar{v}$

$R \neq \sigma^2 I$

- Intuitively, λ_j, denotes the **noise power** along $\bar{\mathbf{u}}_j$.
- Therefore, result says allocate all power to $\bar{\mathbf{u}}_i$...
  - such that **noise power** λ_i is min!!!

choose $\bar{s}$ along direction $\bar{u}_i$ which noise power λ_i is min

$\bar{s} = \bar{\mu}_1 - \bar{\mu}_0 = \bar{u}_i$

To minimize P_D.

Example

- Consider

orthogonal.

$\Rightarrow$ normalize

Rows are orthogonal.

$$\mathbf{R} = \begin{bmatrix} -2 & -2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 \times 2\sqrt{2} \times \sqrt{2} & \\ & 4 \times 2\sqrt{2} \times \sqrt{2} \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$$

Example

- The **eigenvalue decomposition** is given as

$$R = \overbrace{\begin{array}{cc} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \\ \begin{array}{cc} \bar{u}_1 & \quad \bar{u}_2 \end{array} \end{array}}^{U} \underbrace{\begin{bmatrix} \overset{\lambda_1}{8} & \\ & \overset{\lambda_2}{16} \end{bmatrix}}_{\Lambda} \overbrace{\begin{array}{cc} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} & \begin{array}{c} \bar{u}_1^T \\ \bar{u}_2^T \end{array} \end{array}}^{U^T}$$

$$\min_j \lambda_j = \lambda_1 = 8$$

$$\bar{u}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Example

- The **eigenvalue decomposition** is given as

$\bar{u}_1$ λ_1 $\bar{u}_1^T$

$$\begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 16 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$\bar{u}_2^T$

$\bar{u}_2$ λ_2

Example

- Therefore, **eigenvector** corresponding to **minimum eigenvalue** is

$$\bar{\mathbf{u}}_i = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \bar{\boldsymbol{\mu}}_1 - \bar{\boldsymbol{\mu}}_0 = \bar{\mathbf{s}}$$

This is the signal that minimizes the P_e of Gaussian classification

Instructors may use this white area (14.5 cm / 25.4 cm) for the text.
Three options provided below for the font size.

Font: Avenir (Book), Size: 32, Colour: Dark Grey

Font: Avenir (Book), Size: 28, Colour: Dark Grey

Font: Avenir (Book), Size: 24, Colour: Dark Grey

Do not use the space below.