Lecture 11

Karhunen Loeve Representation for non-zero mean random variables / PCA for non-centered data.

$\bar{X}_{nc}$ with $K_{\bar{X}}$ and $\bar{\mu}$.

$\bar{X} = \bar{X}_{nc} - \bar{\mu}$

$\bar{y} = Q^T \bar{X}$ $\quad \bar{X} = Y_1 \bar{q}_1 + Y_2 \bar{q}_2 + \dots Y_n \bar{q}_n$

$Y_i = q_i^T (\bar{X}_{nc} - \bar{\mu})$ $\quad \bar{X}_{nc} = \bar{\mu} + Y_1 \bar{q}_1 + Y_2 \bar{q}_2 + \dots Y_n \bar{q}_n$

$\bar{X}_{app} = \bar{\mu} + Y_1 \bar{q}_1 + \dots Y_M \bar{q}_M$

$\bar{e} = \bar{X}_{nc} - \bar{X}_{app} = Y_{M+1} \bar{q}_{M+1} + \dots Y_n \bar{q}_n$.

$\mathbb{E}\left[\|\bar{e}\|^2\right] = \sum_{i=M+1}^{n} Var(Y_i)$.

Gaussian random vectors:

Last time: If $\bar{X}$ is a Gaussian random vector with invertible covariance matrix $K_{\bar{X}}$ then.

$$f_{\bar{X}}(\bar{x}) = \frac{1}{\sqrt{(2\pi)^n \det(K_{\bar{X}})}} e^{-(\bar{x}-\bar{\mu})^T K_{\bar{X}}^{-1} (\bar{x}-\bar{\mu})/2}.$$

If $K_{\bar{X}}$ is not invertible the density does not exist.

$$K_{\bar{X}} = Q \Lambda Q^T \quad \exists \ \lambda_i = 0.$$

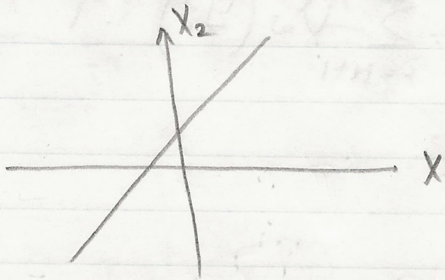
$$\bar{Y} = Q^T \bar{X} \qquad K_{\bar{Y}} = Q^T Q \Lambda Q^T Q = \Lambda$$

$$Var(Y_i) = 0 \ . \quad Y_i = c.$$

$$c = \bar{q}_i^T \bar{X} = q_{i1} X_1 + q_{i2} X_2 + \dots q_{in} X_n .$$

$\Rightarrow$ One of these r.v's can be expressed as a linear function of the others. Equivalently $X_1, X_2 \dots X_n$ take values in a lower dimensional subspace of $\mathbb{R}^n$.

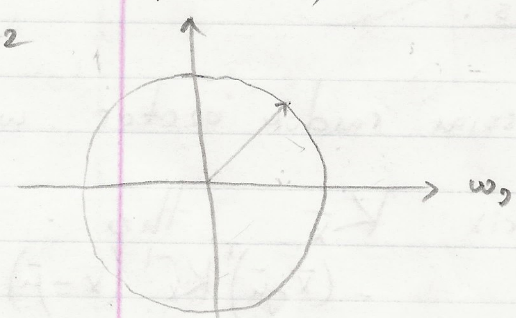
ex $n=2$



Geometry of the Gaussian density

$$\bar{W} \sim N(0, I) \qquad f_{\bar{W}}(\bar{w}) = \frac{1}{(\sqrt{2\pi})^n} e^{-\|\bar{w}\|^2/2} .$$

$n=2$



$f_{\bar{X}}(\bar{x})$ will take the same value for all $\bar{x}$

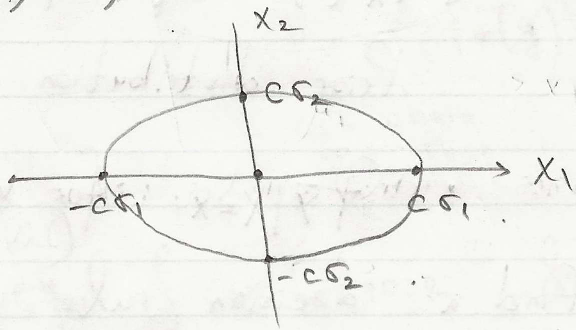
s.t. $(\bar{x}-\bar{\mu})^T K_X^{-1}(\bar{x}-\bar{\mu}) = c^2$

$$K_X = \begin{bmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2\end{bmatrix} \text{ is diagonal}$$

$$(\bar{x}-\bar{\mu})^T \begin{bmatrix}1/\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & 1/\sigma_n^2\end{bmatrix}(\bar{x}-\bar{\mu}) = \sum_{i=1}^{n} (x_i-\mu_i)^2/\sigma_i^2$$

$$\sum_{i=1}^{n} \frac{(x_i-\mu_i)^2}{\sigma_i^2} = c^2.$$

$\underline{n=2}$ $\mu_1=\mu_2=0$. $x_1^2/\sigma_1^2 + \frac{x_2^2}{\sigma_2^2} = c^2$



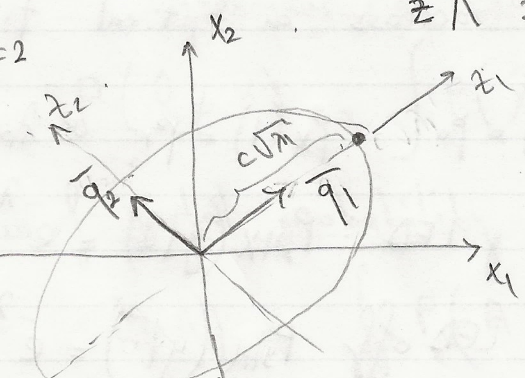
If $K_X = Q\Lambda Q^T$ is not diagonal

$$(\bar{x}-\bar{\mu})^T(Q\Lambda Q^T)^{-1}(\bar{x}-\bar{\mu}) = c^2$$

$$\underbrace{(\bar{x}-\bar{\mu})^T Q}_{\bar{z}^T}\Lambda^{-1}\underbrace{Q^T(\bar{x}-\bar{\mu})}_{\bar{z}} = c^2.$$

$\bar{z}\Lambda^{-1}\bar{z} = c^2$ $\mu_1=\mu_2=0$.

$n=2$



$$\frac{z_1^2}{\lambda_1}+\frac{z_2^2}{\lambda_2} = c^2.$$

- Group meeting
- ~~Yener~~
- Rotation
- ~~Postdoc~~
- ~~DartWarning~~
- Swim

Detection / Hypothesis Testing / Classification

$$\underset{\substack{\uparrow \\ \text{Hypothesis}}}{X} \longrightarrow \underset{\substack{\uparrow \\ \text{Channel}}}{\boxed{f_{Y|X}}} \longrightarrow \underset{\substack{\uparrow \\ \text{observation}}}{\bar{Y}} \longrightarrow \boxed{\text{Decision Rule}} \longrightarrow \hat{X}$$

X : Hypothesis $\quad X \in \{0, 1, \ldots M-1\}$

$\bar{Y}$: Observation $\quad \bar{Y} \in \mathbb{R}^n$

Probabilistic Model: $\cdot\ p_X(x) = P(X = x)$ for $x \in \{0, 1, \ldots M-1\}$

Prior distribution on X.

$\cdot\ f_{\bar{Y}|X=x}$ for $x \in \{0, 1, \ldots M-1\}$
(Likelihoods)

Objective: Find a decision rule $g: \mathbb{R}^n \to \{0, 1, \ldots M-1\}$

$\hat{x} = g(\bar{y})$: guess for X when $\bar{Y} = \bar{y}$

so that $P(X \neq g(\bar{Y})) = P(X \neq \hat{X})$ is minimized

Examples:

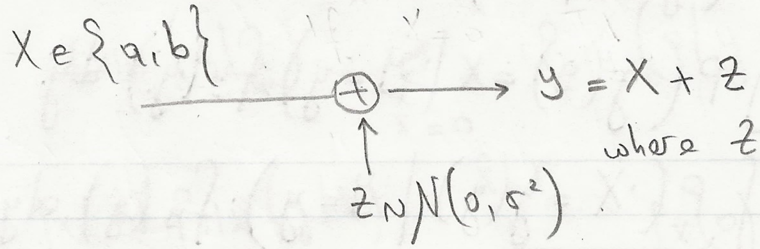
1) Communicate 1-bit accross an optical fiber.

$X \in \{0, 1\}$. $\quad p_X(0) = p_0 \quad p_X(1) = p_1 \quad 0 \leq \lambda_0 \leq \lambda_1$

If $X = 1$, we switch on a LED $\quad p_{y|X}(y|1) = e^{-\lambda_1} \lambda_1^y / y! \quad y \in \mathbb{N}$

If $X = 0$, we switch the LED off $\quad p_{y|H}(y|0) = e^{-\lambda_0} \lambda_0^y / y! \quad y \in \mathbb{N}$

2) Communicate 1-bit over a wireless channel.



where Z is independent of X

<u>Bayes Rule</u>:

if Y is discrete $\quad P_{X|Y}(x|y) = \dfrac{P_X(x)\, P_{Y|X}(y|x)}{P_Y(y)}$. $\qquad P_Y(y) = \sum_x P_X(x) P_{Y|X}(y|x)$

if Y is continious $\quad P_{X|Y}(x|y) = \dfrac{P_X(x)\, f_{Y|X}(y|x)}{f_Y(y)}$

where $\quad f_Y(y) = \sum_x P_X(x) f_{Y|X}(y|x)$

Given $Y = y$, if we choose $\hat{X} = i$, the probability of correct decision is $P_{X|Y}(i|y)$. To maximize the probability of being correct choose $\hat{X}$ s.t.

$$\hat{X} = g_{MAP}(y) = \underset{x \in \{0,1,..M-1\}}{\arg\max}\, P_{X|Y}(x|y).$$

Maximum A Posteriori rule (MAP rule).

$$P_c = P(\hat{X} \neq X) = \int P(\hat{X} \neq X \mid Y = y)\, f_Y(y)\, dy.$$

Law of total probability

$$P(A) = \int P(A \mid Y = y)\, f_Y(y)\, dy.$$

$$
\begin{aligned}
P_c^{(g)} &= P(g(Y) = X) \\
&= \int P(g(Y) = X \mid Y = y)\, f_Y(y)\, dy \\
&= \int \underbrace{P(X = g(y) \mid Y = y)}_{\leq P(X = g_{MAP}(y) \mid Y = y)} f_Y(y)\, dy \\
&\leq \int P(X = g_{MAP}(y) \mid Y = y)\, f_Y(y)\, dy \\
&= P(X = g_{MAP}(Y)) = P_c^{(g_{MAP})}
\end{aligned}
$$

In the special case where $P_X(x) = 1/M$

$$
\begin{aligned}
g_{MAP}(y) &= \underset{x}{\operatorname{argmax}}\; P_{X|Y}(x|y) \\
&= \underset{x}{\operatorname{argmax}}\; \frac{\overset{\frac{1}{M}}{P_X(x)}\, f_{Y|X}(y|x)}{f_Y(y)}
\end{aligned}
$$

doesn't depend on x

$$
= \underset{x}{\operatorname{argmax}}\; f_{Y|X}(y|x)
$$

Maximum Likelihood (ML) rule