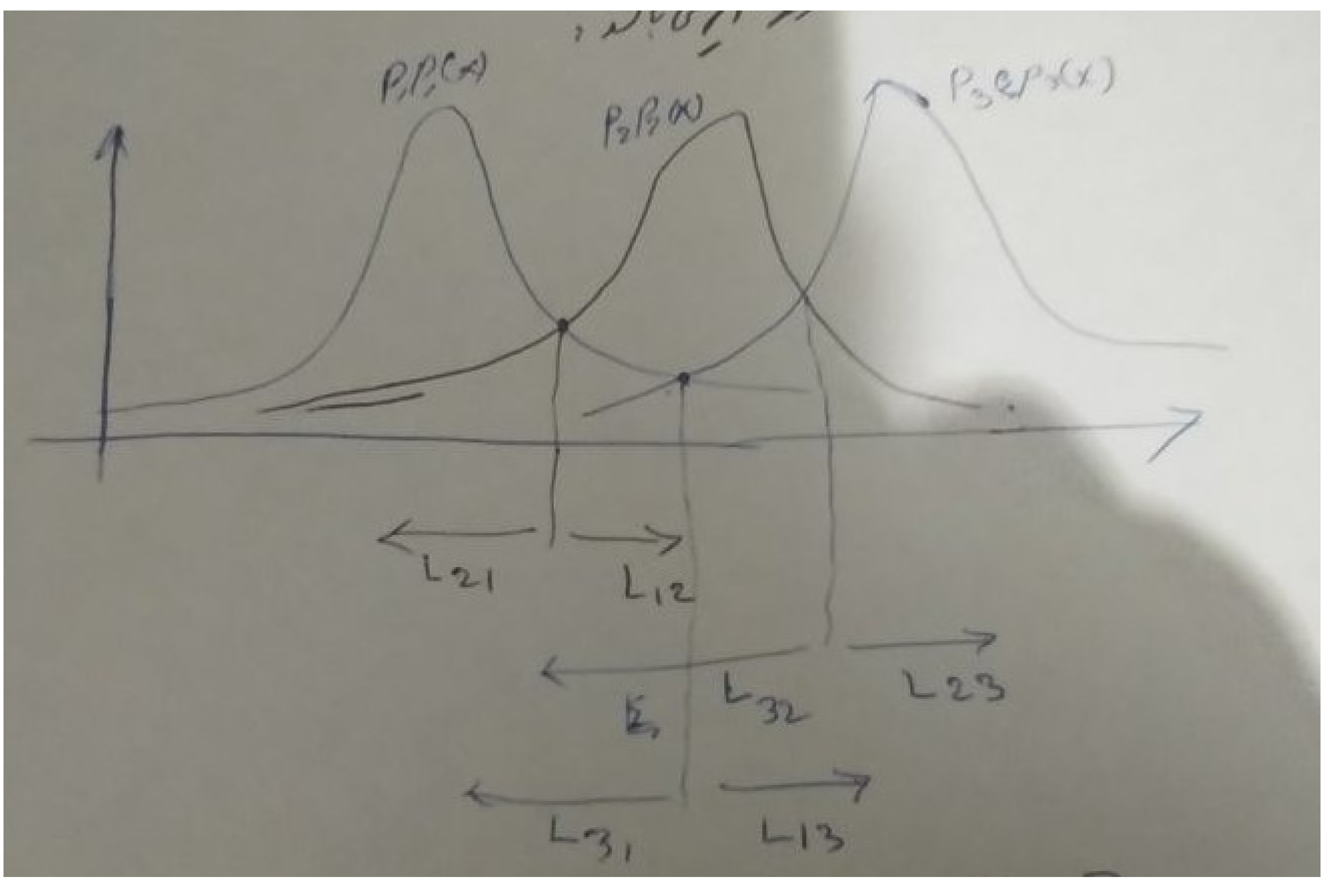
بخش اول تمرین شماره ۱۰ فصل سوم

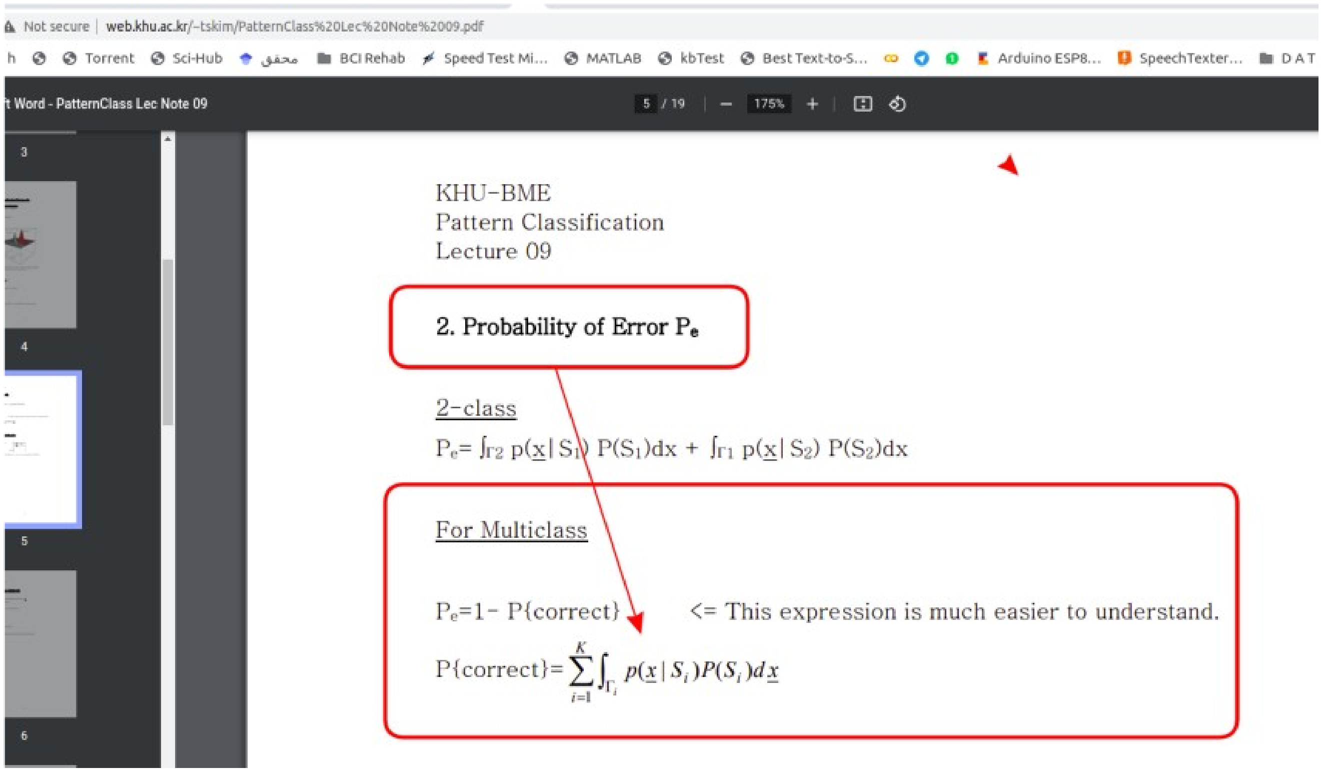
تابع هزینه برای سه وضعیت تصمیم‌گیری W1, W2 و W3، به صورت زیر می باشد

$$R = E[r(x)] = \int Min\,[r_1(x), r_2(x), r_3(x)]\, p(x)\, dx$$

همچنین نمودار خط های تصمیم گیری بر اساس قضیه بیز برای کمترین خطا با فرض توزیع نرمال برای سه کلاس تصمیم‌گیری در تصویر زیر قابل مشاهده می باشد:



بر این اساس و با کمک فرمول زیر:



KHU-BME
Pattern Classification
Lecture 09

2. Probability of Error P_e

2-class

$P_e = \int_{\Gamma_2} p(\underline{x}|S_1)P(S_1)dx + \int_{\Gamma_1} p(\underline{x}|S_2)P(S_2)dx$

For Multiclass

$P_e = 1 - P\{correct\}$ <= This expression is much easier to understand.

$P\{correct\} = \sum_{i=1}^{K}\int_{\Gamma_i} p(\underline{x}|S_i)P(S_i)d\underline{x}$

خطای ریسک تصمیم گیری بیز، برای سه کلاس به صورت زیر می باشد:

$$R_C(w_k|x) = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \begin{bmatrix} P(x|w_1)P(w_1) \\ P(x|w_2)P(w_2) \\ P(x|w_3)P(w_3) \end{bmatrix}$$

Bayes Min Error $\Rightarrow \int \min[R_C(w_k|x)] =$

$$\min \left[\int_{L_{12}} \underbrace{[C_{11}\ C_{12}\ C_{13}]}_{B} \underbrace{[P(x|w_1)P(w_1),\ P(x|w_2)P(w_2),\ P(x|w_3)P(w_3)]^T}_{A} \quad \int_{L_{21}} [C_{21}\ C_{22}\ C_{23}]\cdot A + \int_{L_{23}} [C_{21}\ C_{22}\ C_{23}]\cdot A \quad \int_{L_{32}} [C_{31}\ C_{32}\ C_{33}]\cdot A \right]$$

$$\Rightarrow \mathrm{Min}[R_C] = \begin{cases} \int_{L_{12}} [0 \;\; 1 \;\; 1]\,\underbrace{[N_1(\mu_1,\sigma_1),\, N_2(\mu_2,\sigma_2),\, N_3(\mu_3,\sigma_3)]}_{A}^{T} \\ \int_{L_{21}} [1 \;\; 0 \;\; 0] \cdot [A]^T + \int_{L_{23}} [1 \;\; 0 \;\; 0] \cdot [A]^T \\ \int_{L_{32}} [1 \;\; 0 \;\; 0] \cdot A \end{cases}$$

بر این اساس با استفاده از توابع جدا ساز برای توزیع نرمال بالا شکل ۳ بعدی محدوده های بیز، آن به صورت زیر می باشد:

گزینه های معمول برای توابع جدا ساز:

- تابع ریسک $g_i(x) = -R(\alpha_i \mid x)$
- یا برای حالت کمترین خطا : $g_i(x) = P(\omega_i \mid x)$

در حالتی که $P(x)=1$ (احتمال مشاهده) باشد انتخاب تابع جدا ساز راحت تر می شود: $g_i(x) \equiv p(x \mid \omega_i)\, P(\omega_i)$

چون می توان از طرفین هر تابع یکنوایی گرفت ب اینکه دون نتیجه مقایسه توابع جدا ساز تغییر کند، بخصوص اگر توابع توزیع از نوع نمایی باشد، با گرفتن لگاریتم خواهیم داشت:

$$g_i(x) = \ln p(x \mid \omega_i) + \ln P(\omega_i)$$

توابع جدا ساز برای چگالی توزیع نرمال

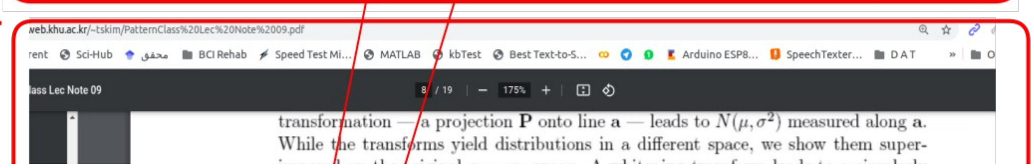
- اگر توزیع کلاس ها نرمال باشد:

$$p(x \mid \omega_i) = \frac{1}{(2\pi)^{d/2} . \Sigma^{1/2}} \exp[\frac{-1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)]$$

- تابع جدا ساز عبارت خواهد بود از:

$$g_i(x) = \ln P(x \mid \omega_i) + \ln P(\omega_i)$$

$$g_i(x) = -\frac{1}{2}(x-\mu_i)\boxed{\sum_i^{-1}}(x-\boxed{\mu_i}) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| + \ln P(\omega_i)$$



transformation — a projection $\mathbf{P}$ onto line $\mathbf{a}$ — leads to $N(\mu, \sigma^2)$ measured along $\mathbf{a}$. While the transforms yield distributions in a different space, we show them superimposed on the original $x_1 - x_2$ space. A whitening transform leads to a circularly symmetric Gaussian, here shown displaced.

<u>Mahalanobis Distance = d_M (DHS p. 36)</u>

$d_M^2(\underline{x},\underline{m})=(\underline{x}-\underline{m})^T\boldsymbol{\Sigma}^{-1}(\underline{x}-\underline{m})$

is the squared Mahalanobis distance from $\underline{\mathbf{x}}$ to $\boldsymbol{\mu}$

The contours of constant density are hyperellipsoids of constant Mahalanobis distance to $\boldsymbol{\mu}$ in Fig. 2.9

$$\boldsymbol{\Sigma}=E\{(\underline{x}-\underline{m})(\underline{x}-\underline{m})^T\}=\begin{bmatrix}\sigma_{11}^2 & \sigma_{12}^2 & \\ \sigma_{21}^2 & & \\ & & \sigma_{NN}^2\end{bmatrix}$$

8

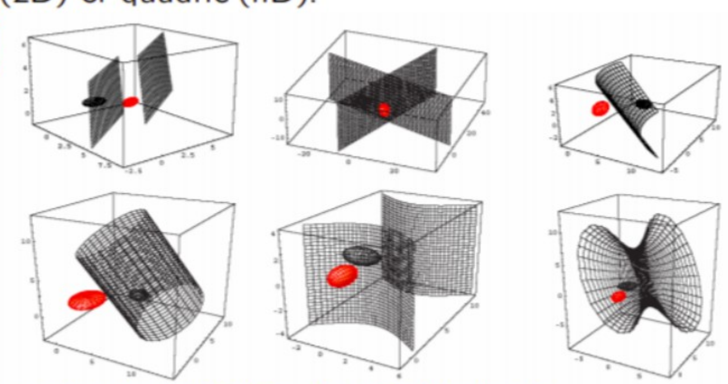
Case 3: Σ_i = arbitrary

The discriminant surface

$$g(\mathbf{x}) = -(\mathbf{x}-\mu_1)^\top\Sigma_1^{-1}(\mathbf{x}-\mu_1)+(\mathbf{x}-\mu_2)^\top\Sigma_2^{-1}(\mathbf{x}-\mu_2)+c'$$

is a conic (2D) or quadric (nD).

e.g. in 3D



The surface can be a hyperboloid, i.e. it need not be closed