Covey, HW 2 Thursday, February 16, 2017 11:37 AM

PROBLEMS

Guilherme França (1)

(1) Consider a two-class classification problem:

$$p(x|w_1) = \mathcal{N}(x|\mu, \Sigma) \quad \text{(class 1)}$$
$$p(x|w_2) = \mathcal{N}(x|-\mu, \Sigma) \quad \text{(class 2)}$$

with $p(w_1) = p(w_2) = \frac{1}{2}$, $\Sigma = \begin{pmatrix} \Sigma_a & 0 \\ 0 & \Sigma_b \end{pmatrix}$, and

$\Sigma_a = \mathrm{diag}(1, 2, \ldots, D)$, $(\Sigma_b)_{ij} = \dfrac{\sqrt{ij}}{2^{|i-j|}} \in \mathbb{R}^{D\times D}$,

μ_a and μ_b are vectors, both in the form $(\underbrace{1, \ldots, 1}_{\text{first } K \text{ components.}}, 0, \ldots, 0)$.

Find the Bayes error L^*.

The decision boundary is given whether the posteriors

$$p(w_1|x) > p(w_2|x) \quad \text{or} \quad p(w_1|x) \leq p(w_2|x).$$

Since $p(w_1) = p(w_2)$ this gives

$$\log p(x|w_1) - \log p(x|w_2) > 0$$
$$-(x-\mu_1)^T\Sigma^{-1}(x-\mu_1) + (x-\mu_2)^T\Sigma^{-1}(x-\mu_2) > 0$$

Because of equal covariance the quadratic terms cancel out and we get

$$2\mu_1^T\Sigma^{-1}x - 2\mu_2^T\Sigma^{-1}x - \mu_1^T\Sigma^{-1}\mu_1 + \mu_2^T\Sigma^{-1}\mu_2 > 0,$$

And since $\mu_2 = -\mu_1 = -\mu$, we have is

$$\boxed{\mu^T\Sigma^{-1}x > 0} \quad \text{(class 1)}$$

or $x \in$ class 2 otherwise.

Thus, the discriminant function is

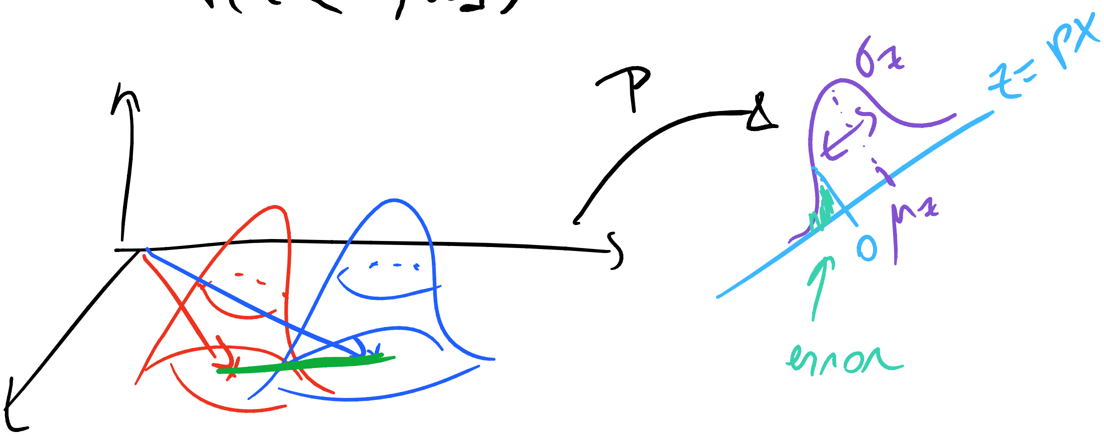
$$g^*(x) = \begin{cases} 1 & \text{if } \mu^T \Sigma^{-1} x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Everything is happening on the line

$$z = PX \,, \quad P \equiv \mu^T \Sigma^{-1}$$

The Bayes error is

$$L^* = P(z > 0, w_2) + P(z \leq 0, w_1)$$
$$= \frac{1}{2} P(z > 0 | w_2) + \frac{1}{2} P(z \leq 0, w_1)$$
$$= P(z \leq 0, w_1)$$



Thus the Bayes error is the "projection" of $N(X | w_1)$ into the line $PX = z$, when $z \leq 0$.
z is a random variable, also normally distributed with

$$\mu_z = E[PX] = PE[X] = P\mu = \mu^T \Sigma^{-1} \mu$$
$$\sigma_z^2 = E[(z - \mu_z)(z - \mu_z)^T]$$
$$= E[P(X - \mu)(X - \mu)^T P^T]$$
$$= P \sigma_X^2 P^T = (\mu^T \Sigma^{-1}) \Sigma (\Sigma^{-1} \mu) = \mu^T \Sigma^{-1} \mu$$

3

Therefore

$$L^* = \int_{-\infty}^{0} dz \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_z} e^{-\frac{1}{2}\frac{(x-\mu_z)^2}{\sigma_z^2}}$$

$$L^* = \int_{-\infty}^{-\frac{\mu_z}{\sigma_z}} dy \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} = \int_{\frac{\mu_z}{\sigma_z}}^{\infty} dy \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$$

Note that $\frac{\mu_z}{\sigma_z} = \frac{\mu^T \Sigma^{-1} \mu}{\sqrt{\mu^T \Sigma^{-1} \mu}} = \sqrt{\mu^T \Sigma^{-1} \mu}$, thus

$$\boxed{L^* = \int_{\sqrt{\mu^T \Sigma^{-1} \mu}}^{\infty} dy \frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}} = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\sqrt{\frac{\mu^T \Sigma^{-1} \mu}{2}}\right)}$$

$$\int_{x}^{\infty} dy \frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}} = \int_0^{\infty} \cdots - \int_0^{x} \cdots$$

$$= \frac{1}{2} - \int_0^{x/\sqrt{2}} \frac{\sqrt{2}\,dt}{\sqrt{2\pi}} e^{-t^2} = \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)$$

$$\operatorname{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

This is the general answer to the problem. Just as a consistency check, if $\Sigma = I$ then $\sqrt{\mu^T \Sigma^{-1} \mu} = \sqrt{\mu^T \mu} = \|\mu\|$, and $L^* = \int_{\|\mu\|}^{\infty} dy \frac{e^{-y^2/2}}{\sqrt{2\pi}}$ which is the error in Trunk's problem.

Now, all we need to do is to specify the above formula to this particular problem.

Since Σ is block diagonal we have

$$\Sigma^{-1} = \begin{pmatrix} \Sigma_a^{-1} & 0 \\ 0 & \Sigma_b^{-1} \end{pmatrix}$$

thus

$$\mu^T \Sigma^{-1} \mu = (\mu_a^T \ \mu_b^T) \begin{pmatrix} \Sigma_a^{-1} & 0 \\ 0 & \Sigma_b^{-1} \end{pmatrix} \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$$

$$= \mu_a^T \Sigma_a^{-1} \mu_a + \mu_b^T \Sigma_b^{-1} \mu_b$$

$$\Sigma_a = \begin{pmatrix} 1 & & \\ & 2 & \\ & & \ddots & \\ & & & D \end{pmatrix} \qquad (\Sigma_b)_{ij} = \frac{\sqrt{ij}}{2^{|i-j|}}$$

diagonal — Not diagonal, but symmetric.

Now if $(\mu_a)_i = \begin{cases} 1 & \text{for } i \leq k \\ 0 & \text{for } i > k \end{cases}$, we then have

$$\mu_a^T \Sigma_a^{-1} \mu_a = \sum_{i=1}^{k} \frac{1}{i} \quad \text{(divergent series)}$$

analogously, $\mu_b = \mu_a$, so

$$\mu_b^T \Sigma_b^{-1} \mu_b = \sum_{i=1}^{k} \sum_{j=1}^{k} (\Sigma_b^{-1})_{ij}$$

Numerically we can easily compute these things. We have

$$f(k) = \mu^T \Sigma^{-1} \mu = \sum_{i=1}^{k} \frac{1}{i} + \sum_{i=1}^{k} \sum_{j=1}^{k} (\Sigma_b)^{-1}_{ij}$$

where $(\Sigma_b)_{ij} = \frac{\sqrt{ij}}{2^{|i-j|}}$. This matrix is invertible.

Therefore,

$$L^*(K) = \frac{1}{2} - \frac{1}{2}\operatorname{erf}\left(\sqrt{\frac{f(K)}{2}}\right).$$

Now consider $\boxed{K=5}$ and any $\boxed{D > K}$.

- Only the first K rows and columns of Σ_b matter:

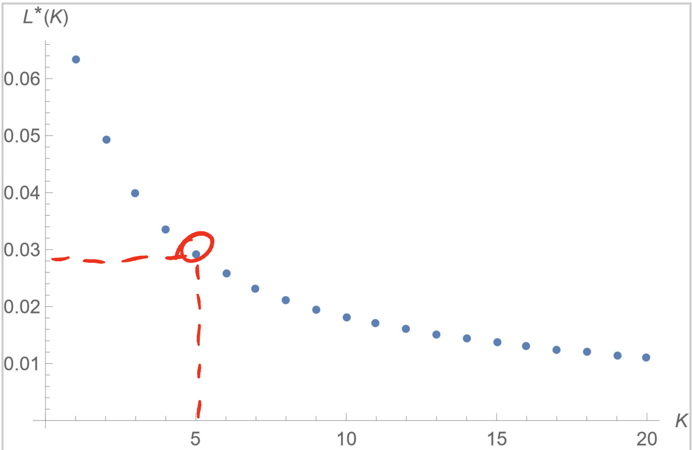
$$f(5) = \sum_{i=1}^{5} \frac{1}{i} + \left(\begin{array}{ccccc|ccc}
4/3 & -\frac{\sqrt{2}}{3} & 0 & 0 & 0 & 0 & \cdots \\
-\frac{\sqrt{2}}{3} & 5/6 & -\frac{\sqrt{2}}{3\sqrt{3}} & 0 & 0 & 0 & \cdots \\
0 & -\frac{\sqrt{2}}{3\sqrt{3}} & 5/9 & -1/\sqrt{3} & 0 & 0 & \cdots \\
0 & 0 & -1/\sqrt{3} & 5/12 & \frac{-1}{3\sqrt{5}} & 0 & \cdots \\
0 & 0 & 0 & \frac{-1}{3\sqrt{5}} & 1/3 & \frac{-\sqrt{2}}{3\sqrt{15}} & \cdots \\
\hline
- & - & - & - & - & 5/18 & \cdots \\
- & - & - & - & - & \vdots & \ddots
\end{array}\right)$$

sum of the elements in the blue block.

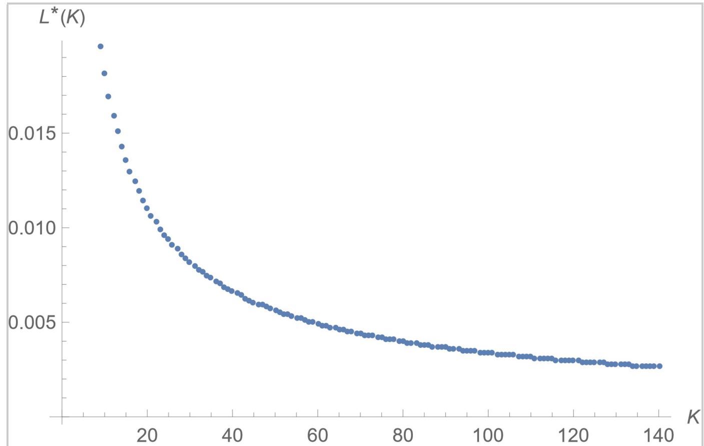
$$= \frac{259}{45} - \frac{2\sqrt{2}}{3\sqrt{3}} - \frac{2\sqrt{2}}{3} - \frac{2}{3\sqrt{3}} - \frac{2}{3\sqrt{5}}$$

Thus $L^* = \frac{1}{2} - \frac{1}{2}\operatorname{erf}\left(\sqrt{\frac{f(5)}{2}}\right) \approx 0.03$.

- This does not change for any $D > K$.
- If K increases, $L^* \to 0$ since the first term in $f(K)$ is a divergent series.
- It may be possible to find Σ_b^{-1} exactly, or at least approximatelly, but this would not change anything since the first term is already divergent. Thus I won't even bother doing that, but rather show the following plots.



(a)



(b)

In the above Fig. we plot $L^*(K)$ versus K, give by the above formula. We fix a large $D \gg K$, thus compute Σ_b^{-1} only once, and vary K. Notice that $L^* \to 0$ as $K \to \infty$ as shown in Fig (b).

(2) Consider the simple hypothesis test

$$H_0 : f = f_A \quad \text{vs} \quad H_A : f = f_A$$

under the rule $f_A(x)/f_0(x) > K \Rightarrow X \sim f_A$, and $f_A(x)/f_0(x) \leq K \Rightarrow X \sim f_0$. This leads to a Hypothesis test with <u>TYPE I</u> error <u>α</u>, and <u>TYPE II</u> error <u>β</u>.

★ Correction!

Now consider two transformations $S: X \mapsto S(X)$ and $T: X \mapsto T(X)$ such that we have

$$\frac{f_A^S(x)}{f_0^S(x)} > K_S \longrightarrow \alpha_S, \beta_S$$

$$\frac{f_A^T(x)}{f_0^T(x)} > K_T \longrightarrow \alpha_T, \beta_T$$

where $\alpha_S = \alpha_T$ and $\beta_S < \beta_T$. Moreover, all these tests are the most powerful.

Show that, if $k_S = k_T$ then there exists priors π $(1-\pi)$ such that $L_S^* < L_T^*$. What can be said if $k_S \neq k_T$?

Let $p_A = \pi$ and $p_0 = 1-\pi$ be the priors. Let's denote $X \in C_0$ if $X \sim f_0$, $X \in C_A$ if $X \sim f_A$. The decision boundary is determined when the posteriors satisfy

$$p(C_A|X) = p(C_0|X)$$

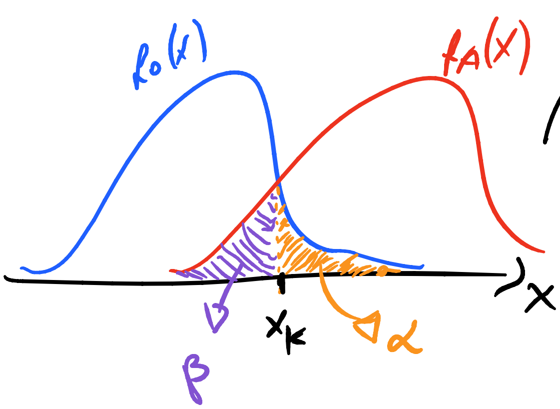
$$\frac{p(X|C_A)\,p_A}{\cancel{p(X)}} = \frac{p(X|C_0)\,p_0}{\cancel{p(X)}}$$

$$f_A(X)\pi = f_0(X)(1-\pi)$$

$$\frac{f_A(X)}{f_0(X)} = \frac{1-\pi}{\pi} \equiv k$$

Thus $\boxed{\pi = \frac{1}{k+1}}$

Notice that



$\boxed{L^* = \pi\beta + (1-\pi)\alpha}$

since the test is MP.

we thus have
$$\begin{cases} L_S^* = \Pi_S \beta_S + (1-\Pi_S)\alpha_S, \\ L_T^* = \Pi_T \beta_T + (1-\Pi_T)\alpha_T \end{cases}$$

Since $\alpha_S = \alpha_T$ we have
$$L_S^* - L_T^* = \Pi_S \beta_S - \Pi_T \beta_T + \alpha_S(\Pi_T - \Pi_S)$$

- If $k_S = k_T$ then $\Pi_S = \Pi_T$, therefore
$$L_S^* - L_T^* = \Pi_S(\beta_S - \beta_T) < 0$$
since $\beta_S < \beta_T$.

- What if $k_S \neq k_T$? Let us consider the two possible cases separatelly.

(a) $k_S > k_T$. Then $\Pi_S = \frac{1}{k_S+1} < \frac{1}{k_T+1} = \Pi_T$. We still have $\beta_S < \beta_T$, thus
$$L_S^* - L_T^* = \underbrace{\Pi_S \beta_S - \Pi_T \beta_T}_{<0} + \alpha_S \underbrace{(\Pi_T - \Pi_S)}_{>0}$$

without further knowledge about $\beta_S, \beta_T, \alpha_S$ the result is inconclusive. Write $\beta_T = \beta_S + \varepsilon$. Then
$$\begin{aligned} L_S^* - L_T^* &= (\Pi_S - \Pi_T)\beta_S - \Pi_T \varepsilon + \alpha_S(\Pi_T - \Pi_S) \\ &= \underbrace{(\Pi_S - \Pi_T)}_{<0}(\beta_S - \alpha_S) \underbrace{- \Pi_T \varepsilon}_{<0} \end{aligned}$$

If $\boxed{\beta_S \geq \alpha_S}$ then certainly $L_S^* < L_T^*$.

If $\boxed{\beta_S < \alpha_S}$ then $L_S^\phi < L_T^\alpha$ only if

$$(\pi_S - \pi_T)(\beta_S - \alpha_S) < \pi_T \varepsilon$$

$$\beta_S - \alpha_S < \frac{\pi_T}{\pi_S - \pi_T} \varepsilon = \frac{1}{\frac{\pi_S}{\pi_T} - 1} \varepsilon$$

$$\alpha_S - \beta_S > \frac{1}{1 - \frac{\pi_S}{\pi_T}} \varepsilon = \frac{1}{1 - \frac{\pi_S}{\pi_T}} (\beta_T - \beta_S)$$

(b) $k_S < K_T$. Then $\pi_S = \frac{1}{k_S + 1} > \frac{1}{k_T + 1} = \pi_T$, thus

$$L_S^\phi - L_T^\alpha = \underbrace{\pi_S \beta_S - \pi_T \beta_T}_{?} + \underbrace{\alpha_S (\pi_T - \pi_S)}_{<0}$$

Again, without further assumptions it is inconclusive.
Write $\beta_T = \beta_S + \varepsilon$ (since $\beta_S < \beta_T$). Then

$$L_S^\phi - L_T^\alpha = \underbrace{(\pi_S - \pi_T)}_{>0}(\beta_S - \alpha_S) - \pi_T \varepsilon$$

If $\boxed{\beta_S \leq \alpha_S}$ then $L_S^\phi < L_T^\alpha$.

If $\boxed{\beta_S > \alpha_S}$ then $L_S^* < L_T^\alpha$ if

$$\pi_T \varepsilon > (\pi_S - \pi_T)(\beta_S - \alpha_S)$$

$$\beta_S - \alpha_S < \frac{\pi_T}{\pi_S - \pi_T} \varepsilon = \frac{\pi_T / \pi_S}{1 - \frac{\pi_T}{\pi_S}} \varepsilon$$

Let us summarize everything.
We have transformations such that

$$\alpha_S = \alpha_T, \quad \beta_S < \beta_T \quad (\beta_T - \beta_S \equiv \varepsilon)$$

β is not power!
power $= 1 - \beta$ in my convention.

Then $L_S^* < L_T^*$ provided:

(1) $k_S = k_T$.

(2) $k_S > k_T$ with $\begin{cases} \beta_S \geq \alpha_S & \text{or} \\ \alpha_S - \beta_S > \dfrac{1}{1 - \frac{\Pi_S}{\Pi_T}}\,\varepsilon & \text{otherwise} \end{cases}$

(3) $k_S < k_T$ with $\begin{cases} \beta_S \leq \alpha_S & \text{or} \\ \beta_S - \alpha_S < \dfrac{\Pi_T}{\Pi_S - \Pi_T}\,\varepsilon & \text{otherwise} \end{cases}$

In any other situation $L_S^* \geq L_T^*$.