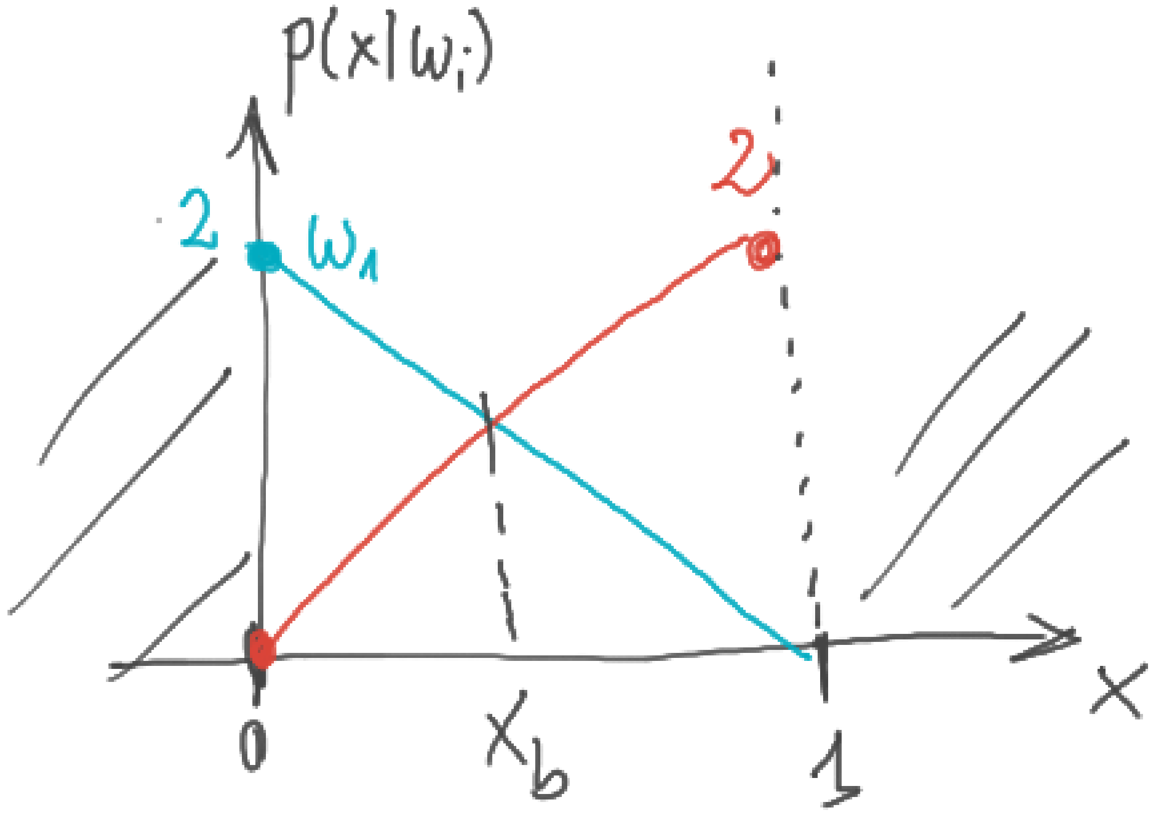
Exercise 1 (10 points)

Let's consider a 2-class problem in a one-dimensional feature space bounded in [0,1], i.e., $x \in [0,1]$. The class-conditional densities are: $p(x|\omega_1) = 2 - 2x$, and $p(x|\omega_2) = 2x$, both defined in [0,1]. Assume that the prior probabilities of the two data classes are $P_1 = P_2$, and that the cost of errors of class ω2 is 1.5 times that of class ω1, that is, $\lambda_{12} = 1.5\lambda_{21}$.

- (5 points) Compute the minimum-risk decision regions, and the Bayesian decision regions, and compute the classification error in both cases.
- (3 points) Plot the joint distributions $P_k p(x|\omega_k)$ for k=1,2 on the one-dimensional feature space, along with the minimum-risk and the Bayesian decision regions.
- (2 points) In the plot, highlight the area(s) corresponding to the additional error incurred when using the minimum-risk decision. One class shows a higher error. *Which one? Why?*



$p(x|\omega_1) = 2 - 2x \qquad P_1 = 1/2$

$p(x|\omega_2) = 2x \qquad P_2 = 1/2$

$$P(\omega_1|x) = P(\omega_2|x)$$

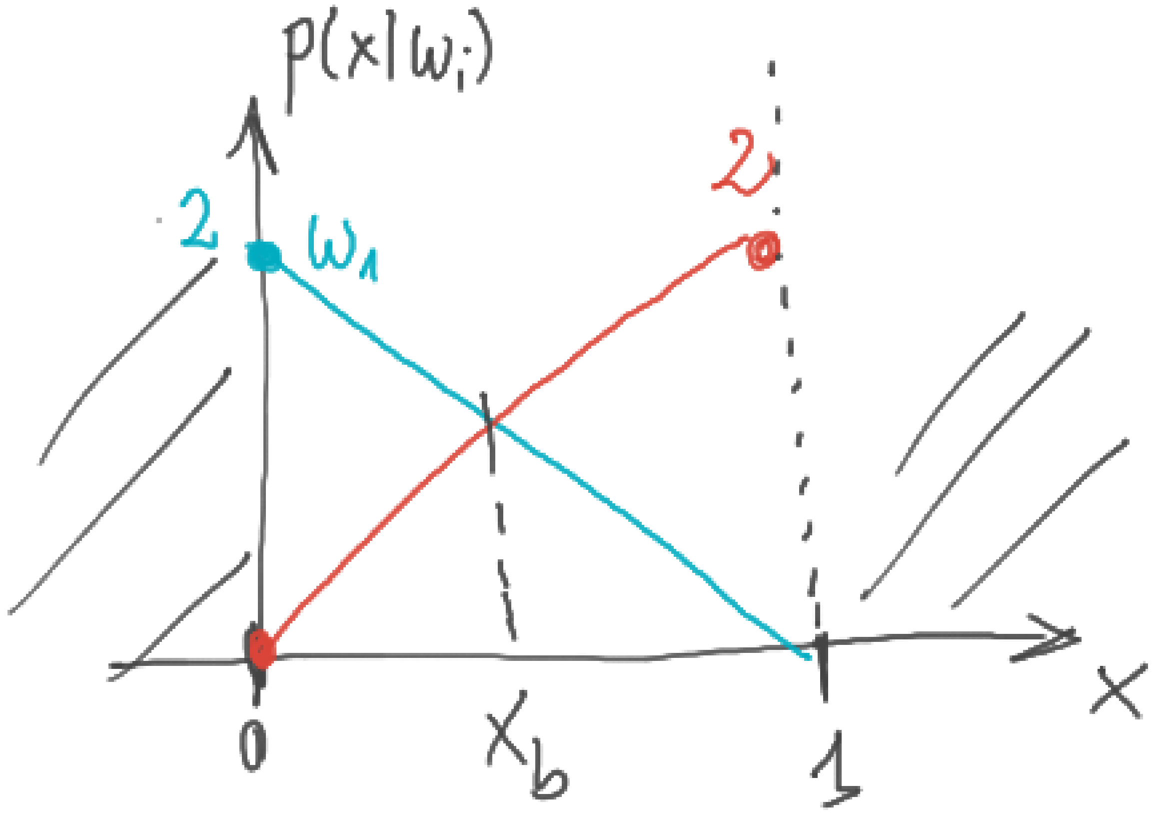
$$\frac{\cancel{P(\omega_1)}\cdot p(x|\omega_1)}{\cancel{p(x)}} = \frac{\cancel{P(\omega_2)}\, p(x|\omega_2)}{\cancel{p(x)}}$$

$$2 - 2x = 2x \qquad 4x = 2 \qquad x_b = 0.5$$

Exercise 1 (10 points)

Let's consider a 2-class problem in a one-dimensional feature space bounded in [0,1], i.e., $x \in [0,1]$. The class-conditional densities are: $p(x|\omega_1) = 2 - 2x$, and $p(x|\omega_2) = 2x$, both defined in [0,1]. Assume that the prior probabilities of the two data classes are $P_1 = P_2$, and that the cost of errors of class ω2 is 1.5 times that of class ω1, that is, $\lambda_{12} = 1.5\lambda_{21}$.

- (5 points) Compute the minimum-risk decision regions, and the Bayesian decision regions, and compute the classification error in both cases.
- (3 points) Plot the joint distributions $P_k p(x|\omega_k)$ for k=1,2 on the one-dimensional feature space, along with the minimum-risk and the Bayesian decision regions.
- (2 points) In the plot, highlight the area(s) corresponding to the <u>additional error</u> incurred when using the minimum-risk decision. One class shows a higher error. *Which one? Why?*



$p(x|\omega_1) = 2 - 2x$ $\quad P_1 = 1/2$

$p(x|\omega_2) = 2x$ $\quad P_2 = 1/2$

$$\Lambda = \begin{array}{cc} & \text{true class} \\ \text{actions} & \begin{bmatrix} 0 & 1.5 \\ 1 & 0 \end{bmatrix} \end{array}$$

λ_{ij} action i but true class is j

$$R_i = \sum_{j=1}^{c} \lambda_{ij} \cdot P(\omega_j|x)$$

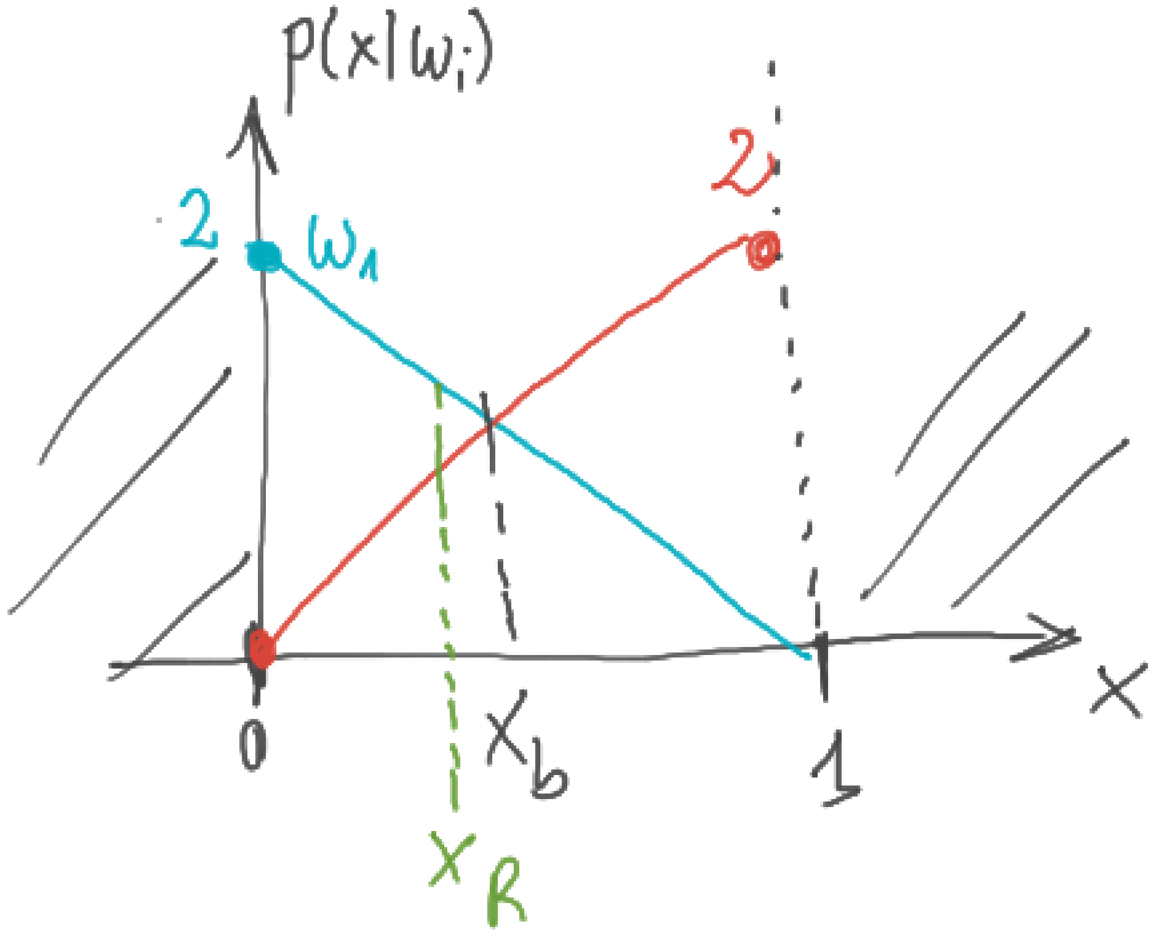
$$R_1 = \cancel{\lambda_{11} \cdot p(x|\omega_1) \cdot P_1} + \lambda_{12}\, p(x|\omega_2) P_2$$

$$R_2 = \lambda_{21} P(x|\omega_1) P_1 + \cancel{\lambda_{22} p(x|\omega_2) P_2}$$

Exercise 1 (10 points)

Let's consider a 2-class problem in a one-dimensional feature space bounded in [0,1], i.e., $x \in [0,1]$. The class-conditional densities are: $p(x|\omega_1) = 2 - 2x$, and $p(x|\omega_2) = 2x$, both defined in [0,1]. Assume that the prior probabilities of the two data classes are $P_1 = P_2$, and that the cost of errors of class ω2 is 1.5 times that of class ω1, that is, $\lambda_{12} = 1.5\lambda_{21}$.

- (5 points) Compute the minimum-risk decision regions, and the Bayesian decision regions, and compute the classification error in both cases.
- (3 points) Plot the joint distributions $P_k p(x|\omega_k)$ for k=1,2 on the one-dimensional feature space, along with the minimum-risk and the Bayesian decision regions.
- (2 points) In the plot, highlight the area(s) corresponding to the additional error incurred when using the minimum-risk decision. One class shows a higher error. *Which one? Why?*



$p(x|\omega_1) = 2 - 2x$ $\quad P_1 = 1/2$

$p(x|\omega_2) = 2x$ $\quad P_2 = 1/2$

$\lambda_{12} = 1.5$

$\lambda_{21} = 1$

$R_1 = \cancel{\lambda_{11} \cdot p(x|\omega_1) \cdot P_1} + \lambda_{12}\, p(x|\omega_2) P_2$

$R_2 = \lambda_{21} P(x|\omega_1) P_1 + \cancel{\lambda_{22}\, p(x|\omega_2) P_2}$

$R_1 = 1.5 \cdot 2x \cdot \frac{1}{2}$

$R_2 = 1.0\,(2 - 2x) \cdot \frac{1}{2}$

$R_1 = R_2$

$1.5 \cdot 2x = 2 - 2x$

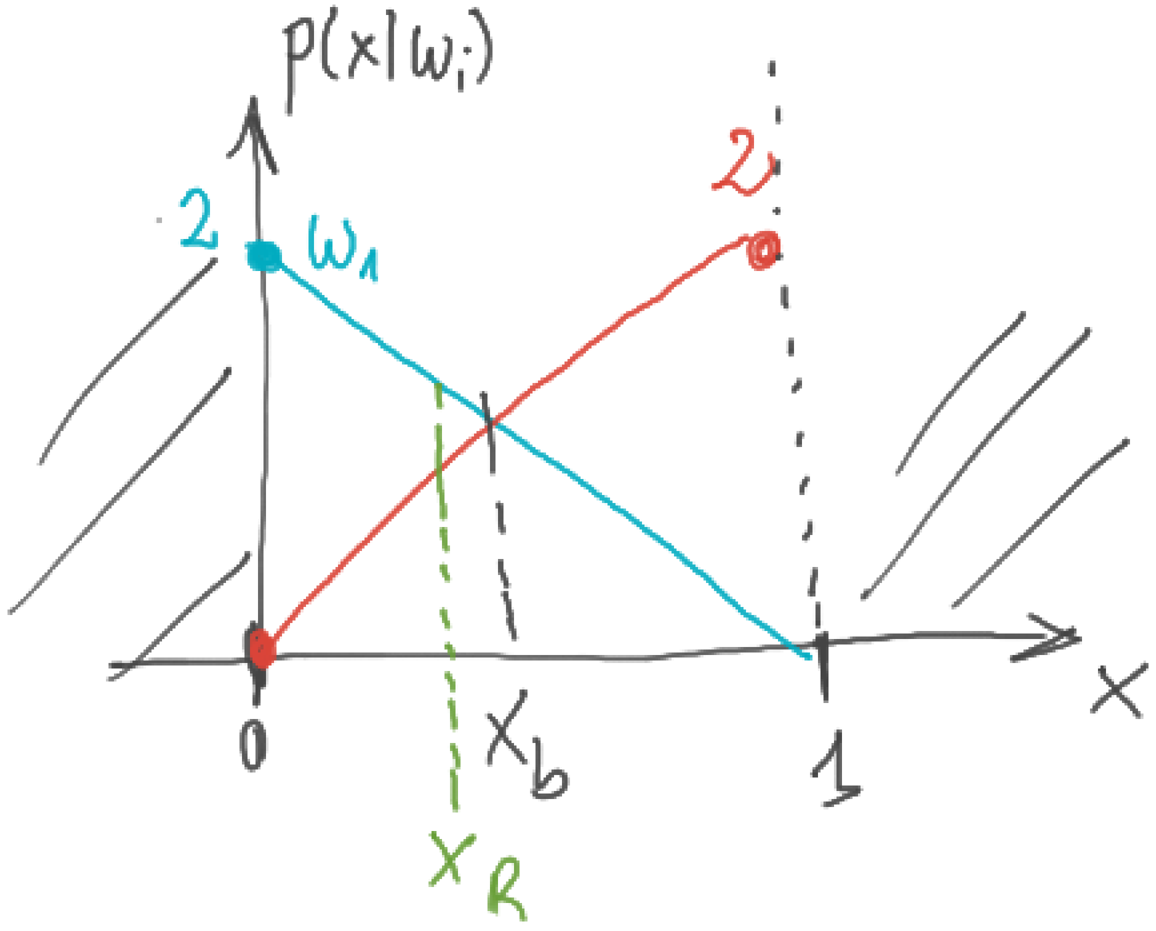
$3x$

$5x = 2$ $\quad x_R = 0.4$

Exercise 1 (10 points)

Let's consider a 2-class problem in a one-dimensional feature space bounded in [0,1], i.e., $x \in [0,1]$. The class-conditional densities are: $p(x|\omega_1) = 2 - 2x$, and $p(x|\omega_2) = 2x$, both defined in [0,1]. Assume that the prior probabilities of the two data classes are $P_1 = P_2$, and that the cost of errors of class ω2 is 1.5 times that of class ω1, that is, $\lambda_{12} = 1.5\lambda_{21}$.

- (5 points) Compute the minimum-risk decision regions, and the Bayesian decision regions, and compute the classification error in both cases.
- (3 points) Plot the joint distributions $P_k p(x|\omega_k)$ for k=1,2 on the one-dimensional feature space, along with the minimum-risk and the Bayesian decision regions.
- (2 points) In the plot, highlight the area(s) corresponding to the additional error incurred when using the minimum-risk decision. One class shows a higher error. *Which one? Why?*



Bayes R₁ $x \in [0, 0.5]$

R₂ $x \in [0.5, 1]$

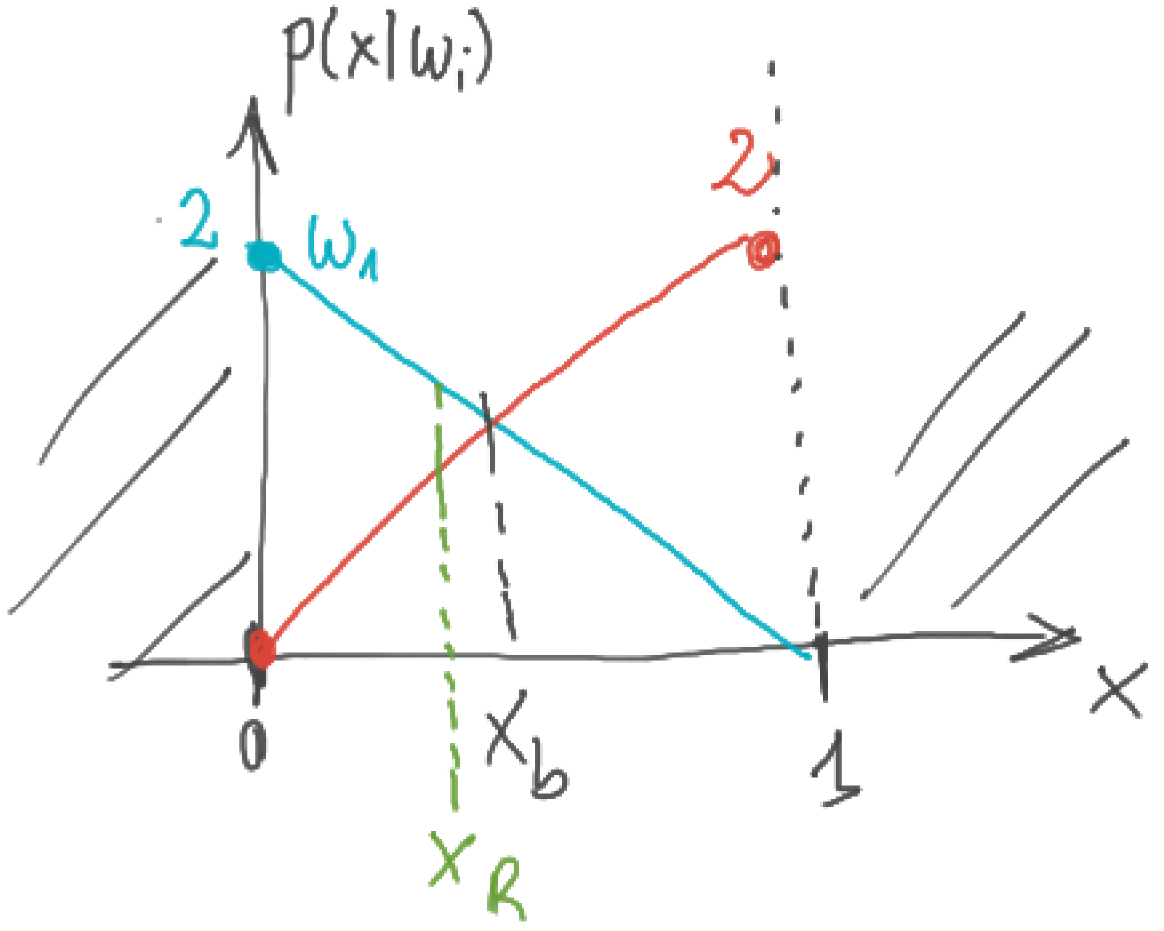
Minimum Risk R₁ $x \in [0, 0.4]$

R₂ $x \in [0.4, 1]$

Exercise 1 (10 points)

Let's consider a 2-class problem in a one-dimensional feature space bounded in [0,1], i.e., $x \in [0,1]$. The class-conditional densities are: $p(x|\omega_1) = 2 - 2x$, and $p(x|\omega_2) = 2x$, both defined in [0,1]. Assume that the prior probabilities of the two data classes are $P_1 = P_2$, and that the cost of errors of class ω2 is 1.5 times that of class ω1, that is, $\lambda_{12} = 1.5\lambda_{21}$.

- (5 points) Compute the minimum-risk decision regions, and the Bayesian decision regions, and compute the classification error in both cases.
- (3 points) Plot the joint distributions $P_k p(x|\omega_k)$ for k=1,2 on the one-dimensional feature space, along with the minimum-risk and the Bayesian decision regions.
- (2 points) In the plot, highlight the area(s) corresponding to the additional error incurred when using the minimum-risk decision. One class shows a higher error. *Which one? Why?*



Bayes $R_1 \quad x \in [0, 0.5]$

$R_2 \quad x \in [0.5, 1]$

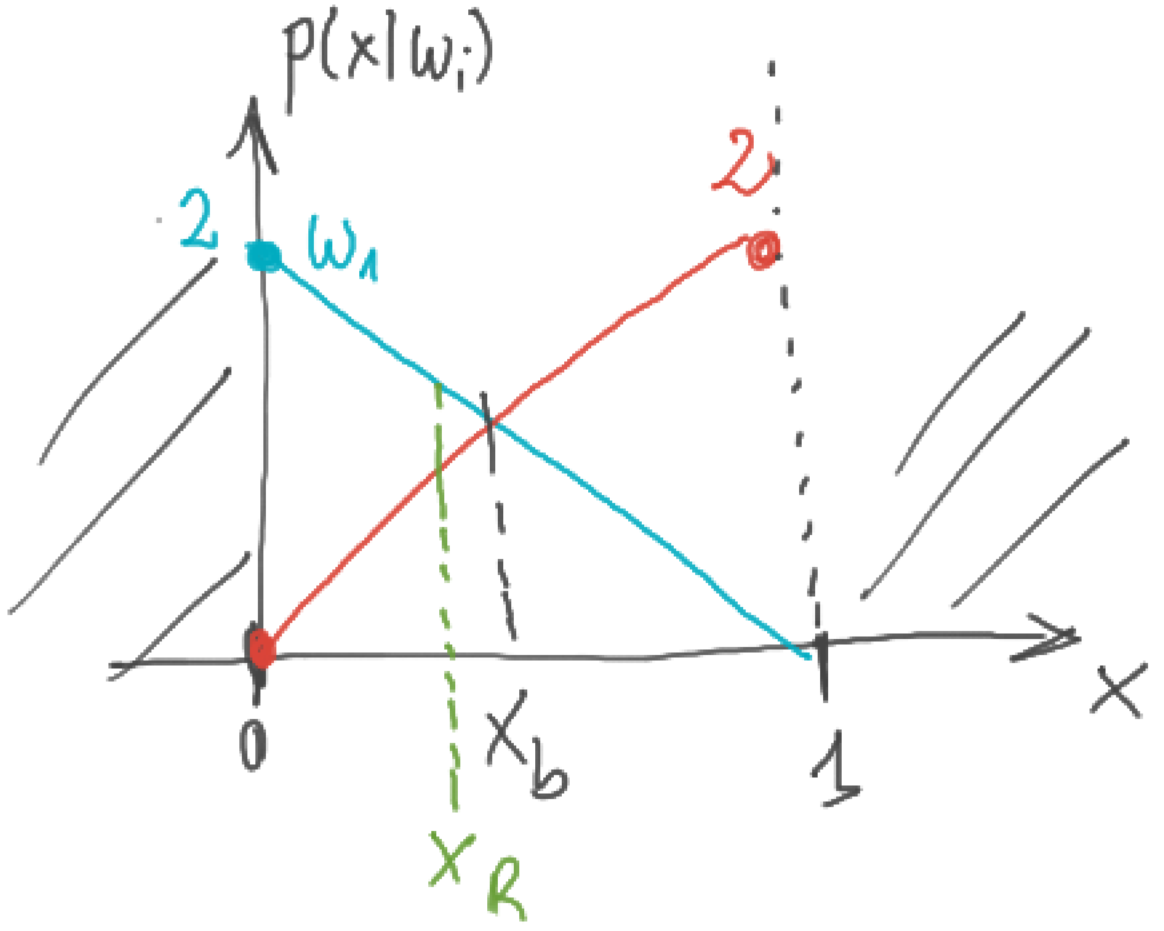
$$P(\text{error}) = P(\text{error}|\omega_1)\cdot P(\omega_1) + P(\text{error}|\omega_2)\cdot P(\omega_2) =$$

$$= P(\omega_1)\cdot \int_{R_2} (2-2x)\,dx + P(\omega_2)\cdot \int_{R_1} (2x)\,dx$$

$$\frac{1}{2}\cdot\left[2x - x^2\right]_{0.5}^{1} + \frac{1}{2}\left[x^2\right]_0^{0.5} = \frac{1}{2}\left[\underbrace{\cancel{2} - \cancel{1} - \cancel{1}}_{1} + 0.25\right] + \frac{1}{2}\left[0.25\right] = 0.25$$

Exercise 1 (10 points)
Let's consider a 2-class problem in a one-dimensional feature space bounded in [0,1], i.e., $x \in [0,1]$. The class-conditional densities are: $p(x|\omega_1) = 2 - 2x$, and $p(x|\omega_2) = 2x$, both defined in [0,1]. Assume that the prior probabilities of the two data classes are $P_1 = P_2$, and that the cost of errors of class ω2 is 1.5 times that of class ω1, that is, $\lambda_{12} = 1.5\lambda_{21}$.

- (5 points) Compute the minimum-risk decision regions, and the Bayesian decision regions, and compute the classification error in both cases.
- (3 points) Plot the joint distributions $P_k p(x|\omega_k)$ for k=1,2 on the one-dimensional feature space, along with the minimum-risk and the Bayesian decision regions.
- (2 points) In the plot, highlight the area(s) corresponding to the additional error incurred when using the minimum-risk decision. One class shows a higher error. *Which one? Why?*



Minimum Risk

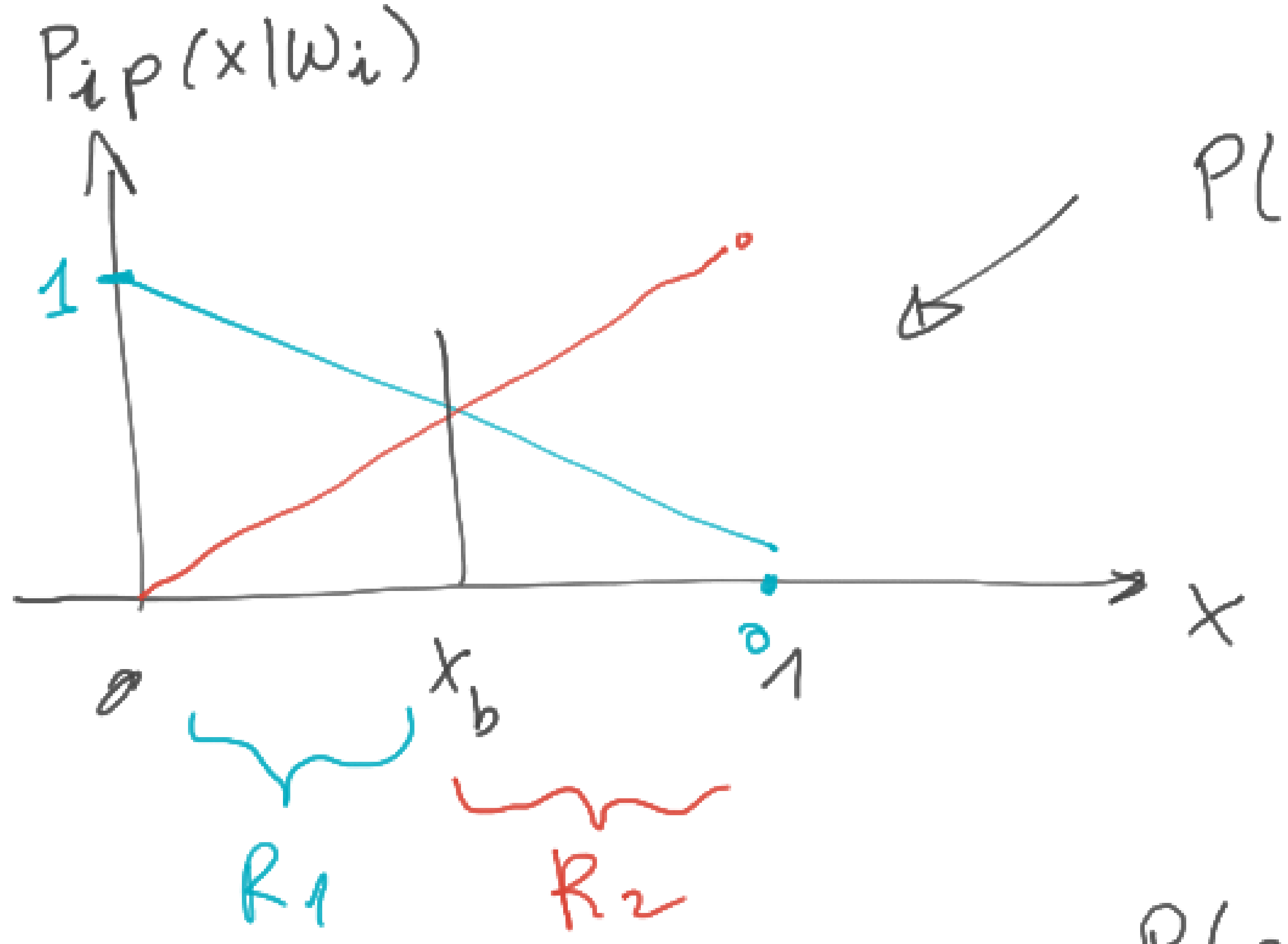
$R_1 \quad x \in [0, 0.4]$

$R_2 \quad x \in [0.4, 1]$

$$P(\text{error}) = P(\text{error}|\omega_1) \cdot P(\omega_1) + P(\text{error}|\omega_2) \cdot P(\omega_2) =$$

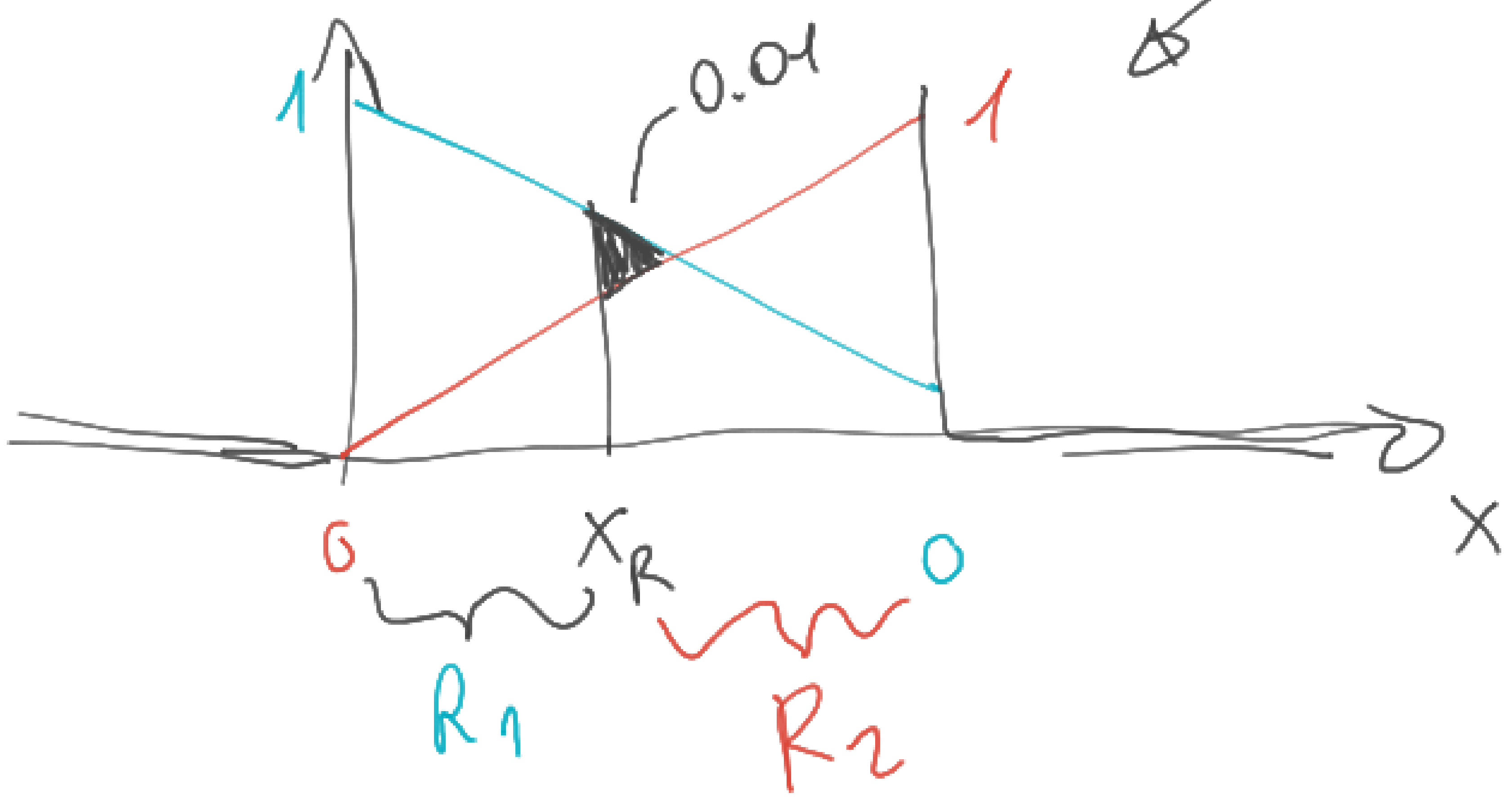
$$= P(\omega_1) \cdot \int_{R_2} (2-2x)\,dx + P(\omega_2) \cdot \int_{R_1} (2x)\,dx$$

$$\frac{1}{2} \cdot \left[2x - x^2\right]_{0.4}^{1} + \frac{1}{2}\left[x^2\right]_0^{0.4} = \frac{1}{2} \cdot \left[\overbrace{2 - 1 - 0.8}^{0.2} + 0.16\right] + \frac{1}{2}\left[0.16\right] = \frac{0.52}{2} = 0.26$$



$P(\text{error}) = 0.25$

Bayes



$P(\text{error}) = 0.26$

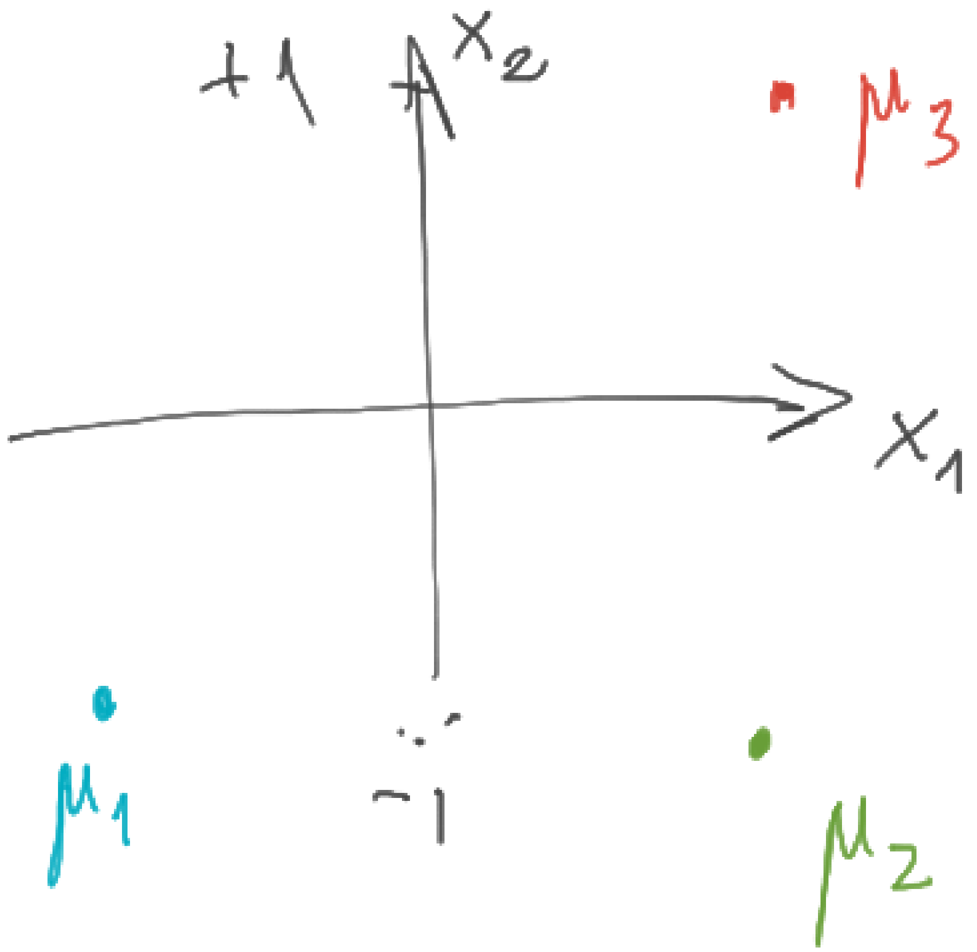
Minimum Risk

Exercise 2 (10 points)

Let us consider a 3-class problem in R^2 (two-dimensional feature space), where the likelihood of each class is Gaussian and given as $p(x|\omega_i) = N(\boldsymbol{\mu}_i, \Sigma_i)$, with

$\Sigma_i = \sigma^2 \mathbf{I}$; $\mu 1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$; $\mu 2 = \begin{pmatrix} +1 \\ -1 \end{pmatrix}$; $\mu 3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$;

and prior probabilities $P_1 = P_2 = P_3$. Compute the decision boundaries and plot them.



$$p(x|\omega_i) = N(\vec{\mu}_i, \vec{\Sigma}_i) \qquad \Sigma_i = \sigma^2 \vec{I}$$

$$\mu_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \qquad \mu_2 = \begin{pmatrix} +1 \\ -1 \end{pmatrix} \qquad \mu_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$P_1 = 1/3 \qquad P_2 = 1/3 \qquad P_3 = 1/3$$

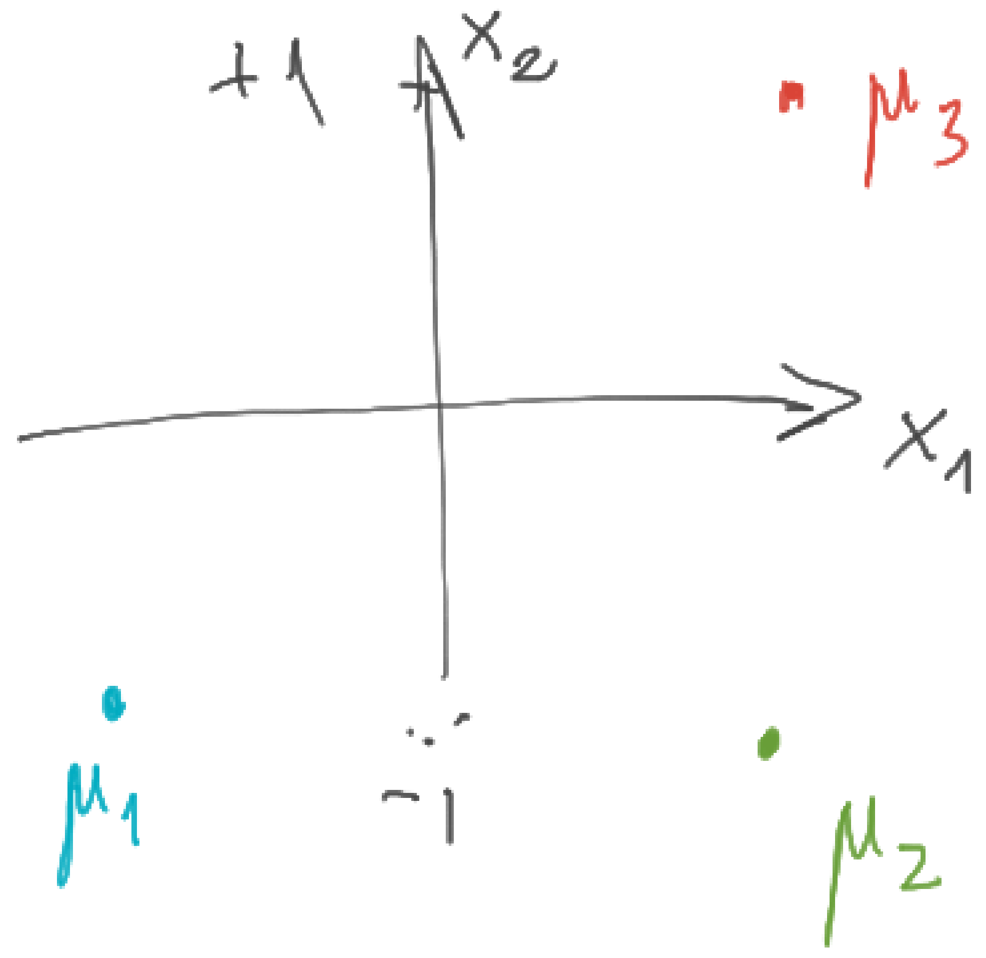
$$p(x|\omega_i) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T \vec{\Sigma}^{-1} (\vec{x}-\vec{\mu})\right]$$

Exercise 2 (10 points)

Let us consider a 3-class problem in R^2 (two-dimensional feature space), where the likelihood of each class is Gaussian and given as $p(x|\omega_i) = N(\boldsymbol{\mu}_i, \Sigma_i)$, with

$\Sigma_i = \sigma^2\mathbf{I}$; $\mu 1 = \begin{pmatrix}-1\\-1\end{pmatrix}$; $\mu 2 = \begin{pmatrix}+1\\-1\end{pmatrix}$; $\mu 3 = \begin{pmatrix}1\\1\end{pmatrix}$;

and prior probabilities $P_1 = P_2 = P_3$. Compute the decision boundaries and plot them.



$$p(x|\omega_i) = N(\vec{\mu}_i, \vec{\vec{\Sigma}}_i) \qquad \Sigma_i = \sigma^2 \vec{I}$$

$$\mu_1 = \begin{pmatrix}-1\\-1\end{pmatrix} \qquad \mu_2 = \begin{pmatrix}+1\\-1\end{pmatrix} \qquad \mu_3 = \begin{pmatrix}1\\1\end{pmatrix}$$

$$P_1 = 1/3 \qquad P_2 = 1/3 \qquad P_3 = 1/3$$

$$g(x) = -\frac{1}{2}(x-\mu)^T(x-\mu) = -\frac{1}{2}\left[\boxed{x^Tx} - 2x^T\mu + \mu^T\mu\right]$$

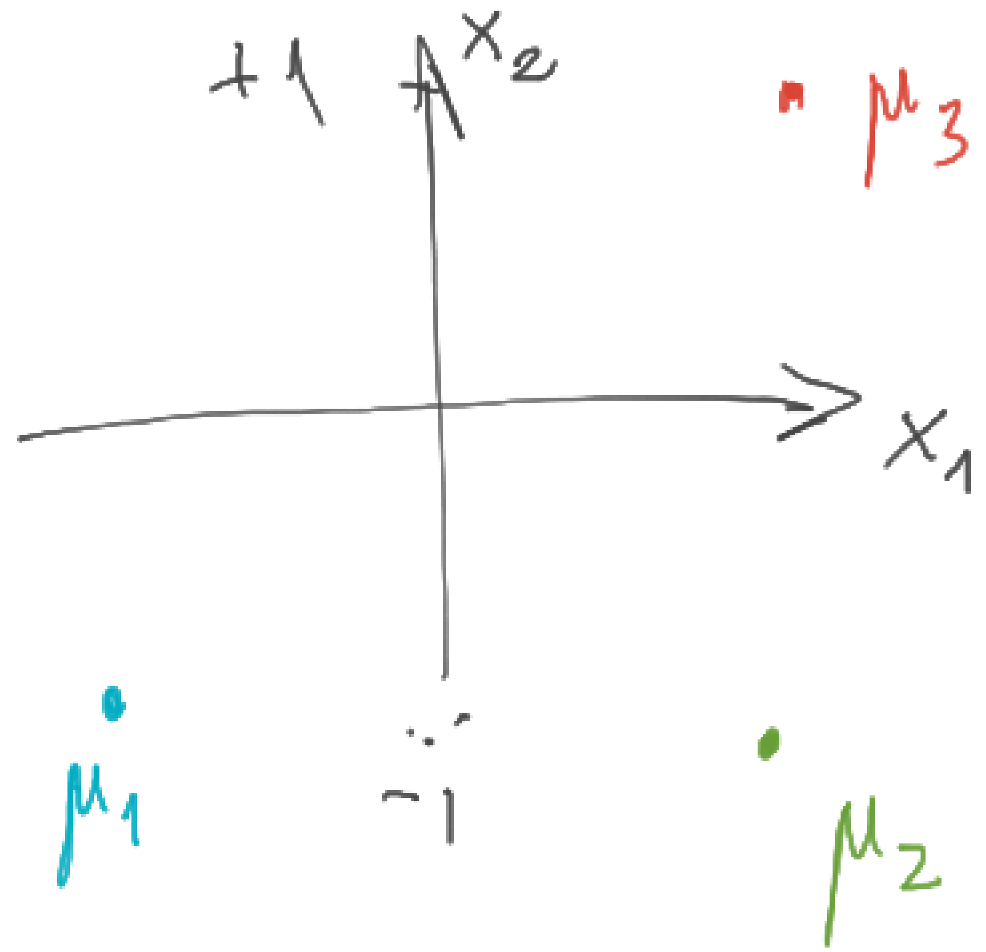
$$g_i(x) = x^T\mu_i - \frac{1}{2}\mu_i^T\mu_i$$

Exercise 2 (10 points)

Let us consider a 3-class problem in R^2 (two-dimensional feature space), where the likelihood of each class is Gaussian and given as $p(x|\omega_i) = N(\boldsymbol{\mu}_i, \Sigma_i)$, with

$\Sigma_i = \sigma^2\mathbf{I}$; $\mu 1 = \begin{pmatrix}-1\\-1\end{pmatrix}$; $\mu 2 = \begin{pmatrix}+1\\-1\end{pmatrix}$; $\mu 3 = \begin{pmatrix}1\\1\end{pmatrix}$;

and prior probabilities $P_1 = P_2 = P_3$. Compute the decision boundaries and plot them.



$$p(x|\omega_i) = N(\vec{\mu}_i, \vec{\vec{\Sigma}}_i) \qquad \Sigma_i = \sigma^2 \vec{I}$$

$$\mu_1 = \begin{pmatrix}-1\\-1\end{pmatrix} \qquad \mu_2 = \begin{pmatrix}+1\\-1\end{pmatrix} \qquad \mu_3 = \begin{pmatrix}1\\1\end{pmatrix}$$

$$g_i(x) = x^T\mu_i - \frac{1}{2}\mu_i^T\mu_i$$

$$\vec{x} = \begin{bmatrix}x_1\\x_2\end{bmatrix} \qquad \mu = [\quad]$$

$$\cancel{-\frac{1}{2}(1+1)} \qquad \cancel{-\frac{1}{2}(1+1)}$$

$$\cancel{-\frac{1}{2}(1+1)}$$

$$g_1(x) = -1x_1 + -1x_2 = -x_1 - x_2$$

$$g_2(x) = x_1 - x_2$$

$$g_3(x) = x_1 + x_2$$

Exercise 2 (10 points)

Let us consider a 3-class problem in R^2 (two-dimensional feature space), where the likelihood of each class is Gaussian and given as $p(x|\omega_i) = N(\boldsymbol{\mu}_i, \Sigma_i)$, with

$\Sigma_i = \sigma^2\mathbf{I}$; $\mu 1 = \begin{pmatrix}-1\\-1\end{pmatrix}$; $\mu 2 = \begin{pmatrix}+1\\-1\end{pmatrix}$; $\mu 3 = \begin{pmatrix}1\\1\end{pmatrix}$;

and prior probabilities $P_1 = P_2 = P_3$. Compute the decision boundaries and plot them.

$$p(x|\omega_i) = N(\vec{\mu}_i, \vec{\vec{\Sigma}}_i)$$

$$g_1(x) = -x_1 - x_2$$

$$g_2(x) = x_1 - x_2$$

$$g_3(x) = x_1 + x_2$$

$B_{12} \rightarrow$ holds for all x where $g_1(x) = g_2(x) > g_3(x)$

$$g_1(x) = g_2(x)$$

$$-x_1 - \cancel{x_2} = x_1 - \cancel{x_2}$$

$$x_1 = 0$$

$$\underline{g_1}(x) > g_3(x)$$

$$-\cancel{x_1} - x_2 > \cancel{x_1} + x_2$$

$$x_2 < 0$$

Exercise 2 (10 points)

Let us consider a 3-class problem in R^2 (two-dimensional feature space), where the likelihood of each class is Gaussian and given as $p(x|\omega_i) = N(\boldsymbol{\mu}_i, \Sigma_i)$, with

$\Sigma_i = \sigma^2 \mathbf{I}$; $\mu 1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$; $\mu 2 = \begin{pmatrix} +1 \\ -1 \end{pmatrix}$; $\mu 3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$;

and prior probabilities $P_1 = P_2 = P_3$. Compute the decision boundaries and plot them.

$$p(x|\omega_i) = N(\vec{\mu}_i, \vec{\vec{\Sigma}}_i)$$

$$g_1(x) = -x_1 - x_2$$

$$g_2(x) = x_1 - x_2$$

$$g_3(x) = x_1 + x_2$$

$B_{23} \rightarrow$

$$g_2(x) = g_3(x)$$

$$\cancel{x_1} - x_2 = \cancel{x_1} + x_2$$

$$x_2 = 0$$

$$g_1(x) < g_2(x) = g_3(x)$$

$$-x_1 - \cancel{x_2} < x_1 - \cancel{x_2}$$

$$x_1 > 0$$

Exercise 2 (10 points)

Let us consider a 3-class problem in R^2 (two-dimensional feature space), where the likelihood of each class is Gaussian and given as $p(x|\omega_i) = N(\boldsymbol{\mu}_i, \Sigma_i)$, with

$\Sigma_i = \sigma^2\mathbf{I}$; $\mu 1 = \begin{pmatrix}-1\\-1\end{pmatrix}$; $\mu 2 = \begin{pmatrix}+1\\-1\end{pmatrix}$; $\mu 3 = \begin{pmatrix}1\\1\end{pmatrix}$;

and prior probabilities $P_1 = P_2 = P_3$. Compute the decision boundaries and plot them.

$$p(x|\omega_i) = N(\vec{\mu}_i, \vec{\vec{\Sigma}}_i)$$

$$g_1(x) = -x_1 - x_2$$

$$g_2(x) = x_1 - x_2$$

$$g_3(x) = x_1 + x_2$$

$B_{13} \rightarrow$

$$g_1(x) = g_3(x)$$

$$-x_1 - x_2 = x_1 + x_2$$

$$x_1 = -x_2$$

$$g_1(x) = g_3(x) > g_2(x)$$

$$-x_1 - x_2 > x_1 - x_2$$

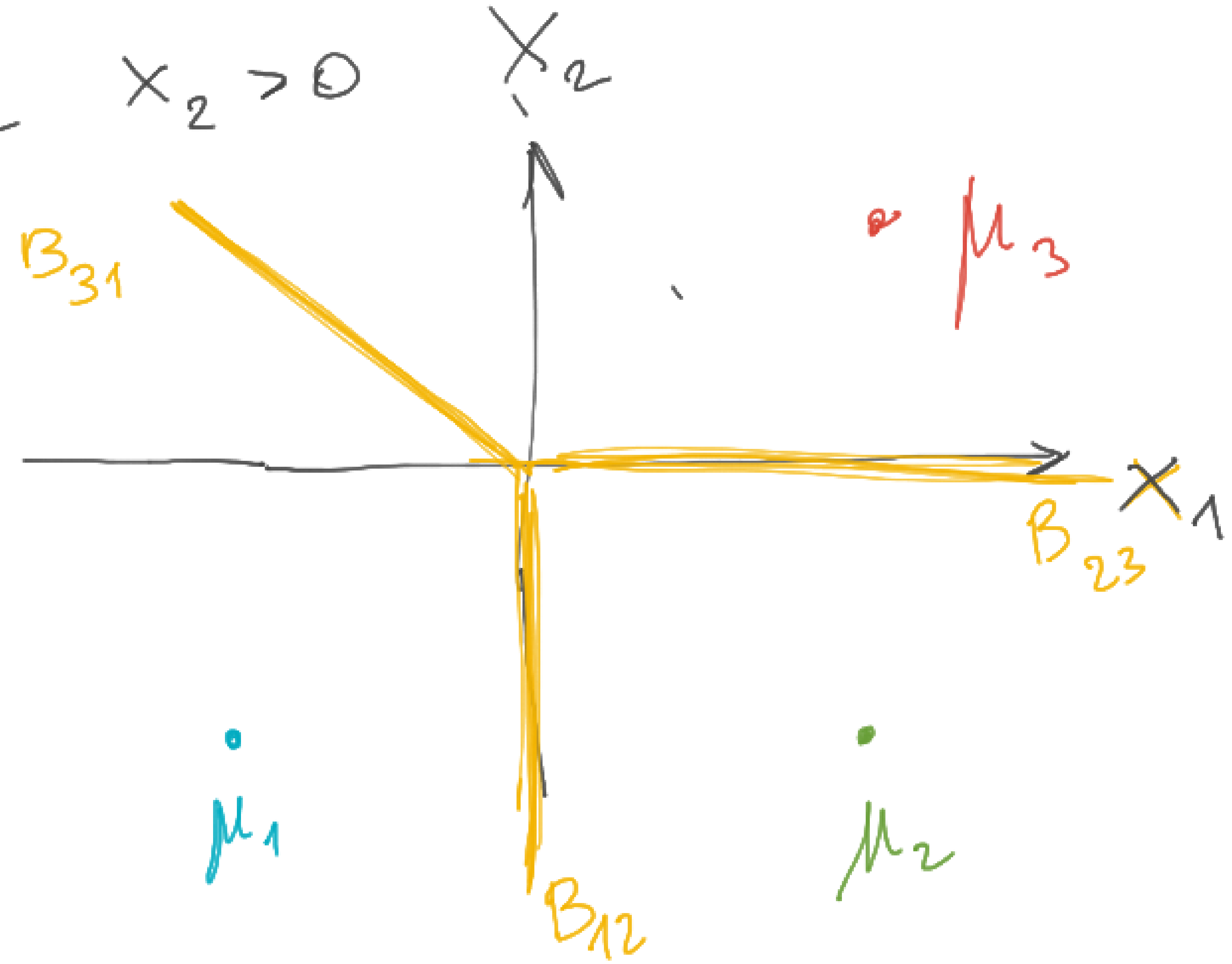
$$\cancel{x_2} - \cancel{x_2} > -x_2 - x_2$$

$$x_2 > 0$$

B_{12} $x_1 = 0$ $x_2 < 0$

B_{23} $x_2 = 0$ $x_1 > 0$

B_{31} $x_1 = -x_2$ $x_2 > 0$



Exercise 3 (10 points)

Given the two-dimensional training points $\mathbf{x_{tr}}$, along with their labels $\mathbf{y_{tr}}$, and a set of test examples $\mathbf{x_{ts}}$, with their labels $\mathbf{y_{ts}}$:

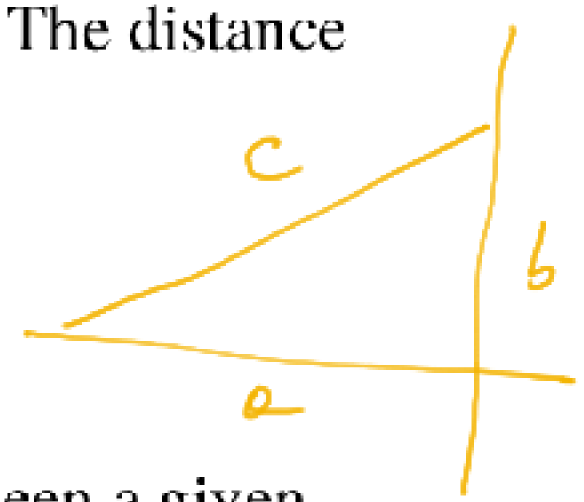
$$\mathbf{x}_{\text{tr}} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 2 & 2 \\ 1 & 0 \\ 3 & 3 \end{bmatrix}, \quad \mathbf{y}_{\text{tr}} = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_{\text{ts}} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{y}_{\text{ts}} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix},$$

classify the points in $\mathbf{x}_{\text{ts}}$ with a k-NN algorithm with k=1, using the l2 distance as the distance metric. The distance matrix computed by comparing $\mathbf{x}_{\text{ts}}$ against $\mathbf{x}_{\text{tr}}$ is given below:

1st col

[[2.24 2.24 1.41 1.00 2.83]
[3.61 3.61 0.00 2.24 1.41]
3° row [1.00 1.00 2.83 1.00 4.24]
[2.83 3.16 1.00 2.00 2.24]]

$c = \sqrt{a^2 + b^2}$



- (5 points) Compute the classification error. In case of equal (minimum) distances between a given test sample and a subset of the training points, assign the test sample to the class of the first point of the training set (from left to right in the distance matrix).
- (5 points) Plot the decision function of the given k-NN classifier.

L_2

$$D = \sqrt{(x_{tr1_1} - x_{ts3_1})^2 + (x_{tr1_2} - x_{ts3_2})^2} = \sqrt{(-1-0)^2 + (0-0)^2} =$$

$$= 1.00$$

Exercise 3 (10 points)

Given the two-dimensional training points $\mathbf{x}_{\mathbf{tr}}$, along with their labels $\mathbf{y}_{\mathbf{tr}}$, and a set of test examples $\mathbf{x}_{\mathrm{ts}}$, with their labels $\mathbf{y}_{\mathrm{ts}}$:

$$\mathbf{x}_{\mathrm{tr}} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 2 & 2 \\ 1 & 0 \\ 3 & 3 \end{bmatrix}, \quad \mathbf{y}_{\mathrm{tr}} = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_{\mathrm{ts}} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{y}_{\mathrm{ts}} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix},$$

classify the points in $\mathbf{x}_{\mathrm{ts}}$ with a k-NN algorithm with k=1, using the l2 distance as the distance metric. The distance matrix computed by comparing $\mathbf{x}_{\mathrm{ts}}$ against $\mathbf{x}_{\mathrm{tr}}$ is given below:

[[2.24 2.24 1.41 1.00 2.83]
 [3.61 3.61 0.00 2.24 1.41]
 [1.00 1.00 2.83 1.00 4.24]
 [2.83 3.16 1.00 2.00 2.24]]

$$y^{pr} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 0 \end{bmatrix} \qquad y_{ts} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix} \qquad \text{Error} = \frac{2}{4} = 0.5 \rightarrow 50\%$$

- (5 points) Compute the classification error. In case of equal (minimum) distances between a given test sample and a subset of the training points, assign the test sample to the class of the first point of the training set (from left to right in the distance matrix).
- (5 points) Plot the decision function of the given k-NN classifier.

Exercise 3 (10 points)

Given the two-dimensional training points $\mathbf{x}_{tr}$, along with their labels $\mathbf{y}_{tr}$, and a set of test examples $\mathbf{x}_{ts}$, with their labels $\mathbf{y}_{ts}$:
$$\mathbf{x}_{tr} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 2 & 2 \\ 1 & 0 \\ 3 & 3 \end{bmatrix}, \quad \mathbf{y}_{tr} = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_{ts} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 0 & 0 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{y}_{ts} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix},$$
classify the points in $\mathbf{x}_{ts}$ with a k-NN algorithm with k=1, using the l2 distance as the distance metric. The distance matrix computed by comparing $\mathbf{x}_{ts}$ against $\mathbf{x}_{tr}$ is given below:

[[2.24 2.24 1.41 1.00 2.83]
[3.61 3.61 0.00 2.24 1.41]
[1.00 1.00 2.83 1.00 4.24]
[2.83 3.16 1.00 2.00 2.24]]

- (5 points) Compute the classification error. In case of equal (minimum) distances between a given test sample and a subset of the training points, assign the test sample to the class of the first point of the training set (from left to right in the distance matrix).
- (5 points) Plot the decision function of the given k-NN classifier.