Homework 2: Bayesian Decision Theory

(DHS) Chapter 2, p65: **2, 6, 37,** and the following.

Problem 4

Two one-dimensional distributions are uniform in [0,2] for ω_1 and [1,4] for ω_2, and P_1=P_2=0.5.

1. Find the Bayes boundary for minimum error, and compute the probability of error.
2. Find the Bayes decision boundary if $\lambda_{11}=\lambda_{22}=0$ and $\lambda_{12}=2\lambda_{21}$.

Problem 5

For a two-class recognition problem with salmon ($\omega = 1$) and sea bass ($\omega = 2$), suppose we have two features $\mathbf{x}$ = (x1, x2) and the two class-conditional densities, $p(x|\omega = 1)$ and $p(x|\omega = 2)$, are 2D Gaussian distributions centered at points (4, 16) and (16, 4) respectively with the same covariance matrix $\mathbf{\Sigma} = 4\mathbf{I}$ (with $\mathbf{I}$ is the identity matrix). Suppose the priors are $P(\omega = 1) = 0.6$ and $P(\omega = 2) = 0.4$.

1. Suppose we use a Bayes decision rule, write the two discriminant functions $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$.
2. Derive the equation for the decision boundary $g_1(\mathbf{x}) = g_2(\mathbf{x})$. Draw the boundary on the feature space (the 2D plane).

Section 2.2

2. Suppose two equally probable one-dimensional densities are of the form $p(x|\omega_i) \propto e^{-|x-a_i|/b_i}$ for $i = 1, 2$ and $0 < b_i$.

(a) Write an analytic expression for each density, that is, normalize each function for arbitrary a_i and positive b_i.

(b) Calculate the likelihood ratio as a function of your four variables.

(c) Sketch a graph of the likelihood ratio $p(x|\omega_1)/p(x|\omega_2)$ for the case $a_1 = 0$, $b_1 = 1$, $a_2 = 1$ and $b_2 = 2$.

6. Consider the Neyman-Pearson criterion for two univariate normal distributions: $p(x|\omega_i) \sim N(\mu_i, \sigma_i^2)$ and $P(\omega_i) = 1/2$ for $i = 1, 2$. Assume a zero-one error loss, and for convenience let $\mu_2 > \mu_1$.

(a) Suppose the maximum acceptable error rate for classifying a pattern that is actually in ω_1 as if it were in ω_2 is E_1. Determine the single-point decision boundary in terms of the variables given.

(b) For this boundary, what is the error rate for classifying ω_2 as ω_1?

(c) What is the overall error rate under zero-one loss?

(d) Apply your results to the specific case $p(x|\omega_1) \sim N(-1, 1)$ and $p(x|\omega_2) \sim N(1, 1)$ and $E_1 = 0.05$.

(e) Compare your result to the Bayes error rate (i.e., without the Neyman-Pearson conditions).

(c) Integrate explicitly the term ...

37. Consider a two-category classification problem in two dimensions with

$$p(\mathbf{x}|\omega_1) \sim N(\mathbf{0}, \mathbf{I}),\ p(\mathbf{x}|\omega_2) \sim N\left(\begin{pmatrix}1\\1\end{pmatrix}, \mathbf{I}\right),\ \text{and } P(\omega_1) = P(\omega_2) = 1/2.$$

(a) Calculate the Bayes decision boundary.

(b) Calculate the Bhattacharyya error bound.

(c) Repeat the above for the same prior probabilities, but

$$p(\mathbf{x}|\omega_1) \sim N\left(\mathbf{0}, \begin{pmatrix}2 & .5\\ .5 & 2\end{pmatrix}\right) \text{ and } p(\mathbf{x}|\omega_2) \sim N\left(\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}5 & 4\\ 4 & 5\end{pmatrix}\right).$$

Nathan Jaggers

Section 2.2

2. Suppose two equally probable one-dimensional densities are of the form $p(x|\omega_i) \propto e^{-|x-a_i|/b_i}$ for $i = 1, 2$ and $0 < b_i$.

(a) Write an analytic expression for each density, that is, normalize each function for arbitrary a_i and positive b_i.

(b) Calculate the likelihood ratio as a function of your four variables.

(c) Sketch a graph of the likelihood ratio $p(x|\omega_1)/p(x|\omega_2)$ for the case $a_1 = 0$, $b_1 = 1$, $a_2 = 1$ and $b_2 = 2$.

a)

$$P(x|\omega_i) \propto e^{-|x-a_i|/b_i}$$

$$P(x|\omega_i) = \frac{e^{-|x-a_i|/b_i}}{\sum_K e^{-|x-a_K|/b_K}}$$

$$P(x|\omega_1) = \frac{e^{-|x-a_1|/b_1}}{\sum_K e^{-|x-a_K|/b_K}} \qquad P(x|\omega_2) = \frac{e^{-|x-a_2|/b_2}}{\sum_K e^{-|x-a_K|/b_K}}$$

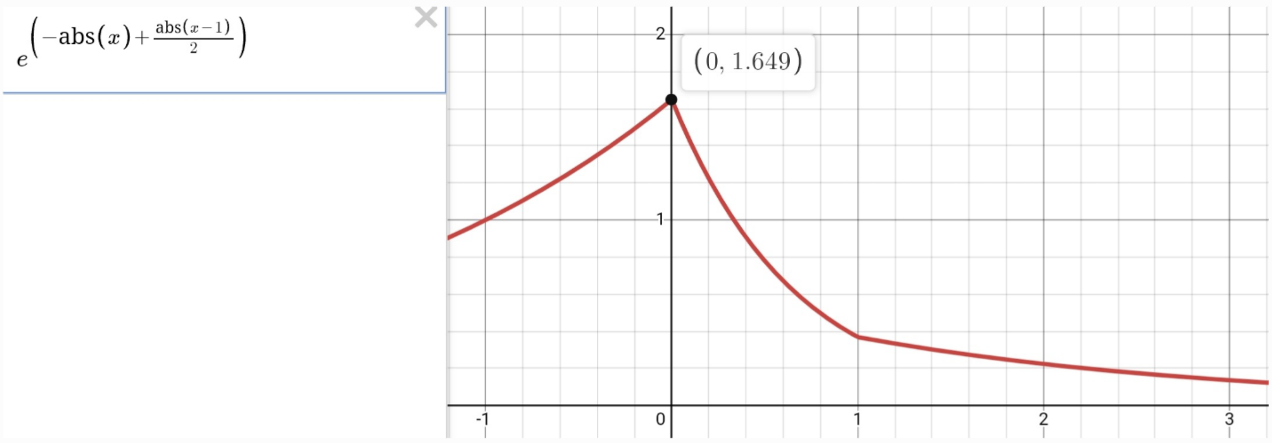
$$\sum = e^{-|x-a_1|/b_1} + e^{-|x-a_2|/b_2}$$

b.)

$$\frac{P(x|w_1)}{P(x|w_2)} = \frac{e^{-|x-a_1|/b_1}}{e^{-|x-a_2|/b_2}} = e^{-\frac{|x-a_1|}{b_1} + \frac{|x-a_2|}{b_2}}$$

c.) $e^{-\frac{|x-a_1|}{b_1} + \frac{|x-a_2|}{b_2}}$ & $a_1 = 0$, $b_1 = 1$, $a_2 = 1$, $b_2 = 2$

$$e^{\left(-|x| + \frac{|x-1|}{2}\right)}$$



6. Consider the Neyman-Pearson criterion for two univariate normal distributions: $p(x|\omega_i) \sim N(\mu_i, \sigma_i^2)$ and $P(\omega_i) = 1/2$ for $i = 1, 2$. Assume a zero-one error loss, and for convenience let $\mu_2 > \mu_1$.

(a) Suppose the maximum acceptable error rate for classifying a pattern that is actually in ω_1 as if it were in ω_2 is E_1. Determine the single-point decision boundary in terms of the variables given.

(b) For this boundary, what is the error rate for classifying ω_2 as ω_1?

(c) What is the overall error rate under zero-one loss?

(d) Apply your results to the specific case $p(x|\omega_1) \sim N(-1, 1)$ and $p(x|\omega_2) \sim N(1, 1)$ and $E_1 = 0.05$.

(e) Compare your result to the Bayes error rate (i.e., without the Neyman-Pearson conditions).

a)

$$P(\text{error}|x) = \begin{cases} P(\omega_1|x) & \text{if we decide } \omega_2 \quad < E_1 \\ P(\omega_2|x) & \text{if we decide } \omega_1 \end{cases}$$

$$P(\text{error}) = \int_{-\infty}^{\infty} P(\text{error}|x)\, p(x)\, dx$$

$$P(\text{error}_{\omega_1 \text{ as } \omega_2}) = \int_{t}^{\infty} P(\omega_1|x)\, P(x)\, dx$$

where t is boundary

$$= \int_{t}^{\infty} p(x|\omega_1)\, P(\omega_1)\, dx$$

$$= \int_{t}^{\infty} (N(\mu_1, \omega_1))\left(\tfrac{1}{2}\right) dx \quad < E_1$$

$$= \frac{1}{2} \int_{t}^{\infty} N(\mu_1, \omega_1)\, dx \quad < E_1$$

we could evaluate the integral & then solve for t to give our boundary

$$= \frac{1}{2}\int_{t}^{\infty} N(\mu_1, \sigma_1^2)\, dx \leq E_1$$

Not an elementary integral

we could evaluate the integral & then solve for t to give our boundary based on E_1

b)

$$\int_{-\infty}^{t} p(x|\omega_2) P(\omega_2)\, dx$$

$$= \frac{1}{2}\int_{-\infty}^{t} N(\mu_2, \sigma_2^2)\, dx \leq E_2$$

c)

Overall error is the sum of two previous errors

$$E_{total} = E_1 + E_2$$

d) $$\frac{1}{2}\int_{t}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(-0.5(x-\mu)^2/\sigma^2)} = E_1$$

$N(\mu_1, \sigma_1^2) \rightarrow N(-1, 1)$, $E_1 = 0.05$

used Matlab & online resources to find boundary

$t = 0.2815$

$$\frac{1}{2}\int_{-\infty}^{0.2815} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(-0.5(x-\mu)^2/\sigma^2)} = E_2$$

$N(\mu_2, \sigma_2^2) \rightarrow N(1, 1)$

$E_2 = 0.118112$

$E_{total} = 0.168117$

e) means are set @ -1 & 1 with same cov, therefore min error boundary likely at 0. distributions are identical otherwise so we can simplify math

$$E_{B\,total} = 2\int_0^{\infty} \frac{1}{2} N(\mu_1, \sigma_1^2)\, dx$$

$$= \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{(-0.5(x-\mu_1)^2/\sigma_1^2)}$$

$$= \int_0^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{(-0.5(x+1)^2)}$$

Non-elementary, using matlab

$$E_{B\,total} = 0.158655$$

---

a) $\frac{1}{2}\int_{-\infty}^{t} N(\mu_2, \sigma_2^2)\, dx \leq E_2$

b) $\frac{1}{2}\int_{-\infty}^{t} N(\mu_2, \sigma_2^2)\, dx \leq E_2$

c) $E_{Total} = E_1 + E_2$

d) $t = 0.2815$ $\quad E_2 = 0.118112$ $\quad E_{Total} = 0.168117$

e) $E_{B\,total} = 0.158655$

(c) Integrate explicitly ...

37. Consider a two-category classification problem in two dimensions with

$$p(\mathbf{x}|\omega_1) \sim N(\mathbf{0}, \mathbf{I}),\ p(\mathbf{x}|\omega_2) \sim N\left(\begin{pmatrix}1\\1\end{pmatrix}, \mathbf{I}\right),\ \text{and } P(\omega_1) = P(\omega_2) = 1/2.$$

(a) Calculate the Bayes decision boundary.

(b) Calculate the Bhattacharyya error bound.

(c) Repeat the above for the same prior probabilities, but

$$p(\mathbf{x}|\omega_1) \sim N\left(\mathbf{0}, \begin{pmatrix}2 & .5\\ .5 & 2\end{pmatrix}\right) \text{ and } p(\mathbf{x}|\omega_2) \sim N\left(\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}5 & 4\\ 4 & 5\end{pmatrix}\right).$$

$P(\hat{x}|\omega_1) \sim \hat{mu} = \begin{pmatrix}0\\0\end{pmatrix}$ $\hat{\Sigma} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$

$P(\omega_1) = 1/2$

$P(\hat{x}|\omega_2) \sim \hat{mu} = \begin{pmatrix}1\\1\end{pmatrix}$ $\hat{\Sigma} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$

$P(\omega_2) = 1/2$

$$\mathbf{w}^T(\mathbf{x} - \mathbf{x}_0) = 0$$
$$\mathbf{w} = \boldsymbol{\mu}_i - \boldsymbol{\mu}_j$$
$$\mathbf{x}_0 = \frac{1}{2}(\boldsymbol{\mu}_i + \boldsymbol{\mu}_j) - \frac{\sigma^2}{||\boldsymbol{\mu}_i - \boldsymbol{\mu}_j||^2} ln \frac{P(\omega_i)}{P(\omega_j)} (\boldsymbol{\mu}_i - \boldsymbol{\mu}_j)$$

$$W = \begin{pmatrix}0\\0\end{pmatrix} - \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}-1\\-1\end{pmatrix}$$

$$X_0 = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix} - \cancel{\frac{1}{1} \ln\left(\frac{1/2}{1/2}\right)\begin{pmatrix}1\\1\end{pmatrix}} \rightarrow 0$$

$$X_0 = \begin{pmatrix}1/2\\1/2\end{pmatrix}$$

$$\begin{pmatrix}-1 & -1\end{pmatrix}\begin{pmatrix}x_1 - 1/2\\ x_2 - 1/2\end{pmatrix} = 0$$

$$-X_1 + 1/2 - X_2 + 1/2 = 0$$

$$X_1 + X_2 = 1 \quad \text{or} \quad X_2 = -X_1 + 1$$

b)

$$P(error) \le \sqrt{P(\omega_1)P(\omega_2)}e^{-k\left(\frac{1}{2}\right)}$$

$$k\left(\frac{1}{2}\right) = \frac{1}{8}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T \left[\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}\right]^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) + \frac{1}{2} ln \frac{\left|\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}\right|}{\sqrt{|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|}}$$

$\det(\Sigma_1) = 1$ & Σ_2

$\begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}0\\0\end{pmatrix}$ $\quad \begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}/2$ $\quad \begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}0\\0\end{pmatrix}$ $\quad \begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}/2$ $\quad \sqrt{1 \cdot 1} = 1$

$$\frac{1}{8}\begin{pmatrix}1 & 1\end{pmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{pmatrix}1\\1\end{pmatrix} + \frac{1}{2}\ln\left(\frac{1}{1}\right)$$

$$\begin{pmatrix}\underset{1}{1+0} & \underset{1}{0+1}\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}$$

$$\frac{1}{8} \cdot 2 = 1/4$$

$$P(error) \le \sqrt{P(\omega_1)P(\omega_2)}\; e^{-(1/4)}$$

$$= \sqrt{1/2 \cdot 1/2} \cdot e^{-1/4}$$

$$1/2 \cdot 0.7788 = 0.0389$$

c)

$P(\hat{x} | \omega_1) \sim \quad \hat{mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \hat{\Sigma} = \begin{bmatrix} 2 & 1/2 \\ 1/2 & 2 \end{bmatrix}$

$P(\omega_1) = 1/2$

$P(\hat{x} | \omega_2) \sim \quad \hat{mu} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \hat{\Sigma} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$

$P(\omega_2) = 1/2$

$$g_i(\mathbf{x}) = \mathbf{x}^t \mathbf{W}_i \mathbf{x} + \boldsymbol{w_i}^t \mathbf{x} + w_{i0}$$

(quadratic discriminant)

where $\mathbf{W}_i = -\frac{1}{2}\boldsymbol{\Sigma}_i^{-1}$, $\boldsymbol{w_i} = \boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i$, and

$w_{i0} = -\frac{1}{2}\boldsymbol{\mu}_i^t \boldsymbol{\Sigma}_i^{-1}\boldsymbol{\mu}_i - \frac{1}{2}\ln|\boldsymbol{\Sigma}_i| + \ln P(\omega_i)$

$\Sigma_i^{-1} = \begin{bmatrix} 8/15 & -2/15 \\ -2/15 & 8/15 \end{bmatrix}$

$W_i = \begin{bmatrix} -4/15 & 1/15 \\ 1/15 & -4/15 \end{bmatrix}$

$w_i = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$w_{i0} = -1.3540$

$\Sigma_j^{-1} = \begin{bmatrix} 5/9 & -4/9 \\ -4/9 & 5/9 \end{bmatrix}$

$W_j = \begin{bmatrix} -5/18 & 2/9 \\ 2/9 & -5/18 \end{bmatrix}$

$w_j = \begin{bmatrix} 1/9 \\ 1/9 \end{bmatrix}$

$w_{j0} = -1.9029$

$g_i = -\frac{4}{15}x_2^2 + x_2 x_1 \frac{1}{15} - \frac{4}{15}x_1^2 + x_2 x_1 \frac{1}{15} - 1.354$

$-\frac{4}{15}x_2^2 + \frac{2}{15}x_2 x_1 - \frac{4}{15}x_1^2 - 1.354$

$g_j = -\frac{5}{18}x_1^2 + \frac{8}{18}x_1 x_2 - \frac{5}{18}x_2^2 + \frac{1}{9}x_1 + \frac{1}{9}x_2 - \frac{1}{9} - 1.1 + \ln\frac{1}{2}$

$0 = g_i - g_j = x_1^2 + x_2^2 - 28x_1x_2 - 10x_1 - 10x_2 + 50$

$$P(error) \le \sqrt{P(\omega_1)P(\omega_2)}e^{-k\left(\frac{1}{2}\right)}$$

$$k\left(\frac{1}{2}\right) = \frac{1}{8}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T \left[\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}\right]^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) + \frac{1}{2} ln \frac{\left|\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}\right|}{\sqrt{|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|}}$$

using matlab to calculate

$k(1/2) = 0.1499$

$$P(error) \le \sqrt{P(\omega_1)P(\omega_2)}\, e^{-(0.1499)}$$

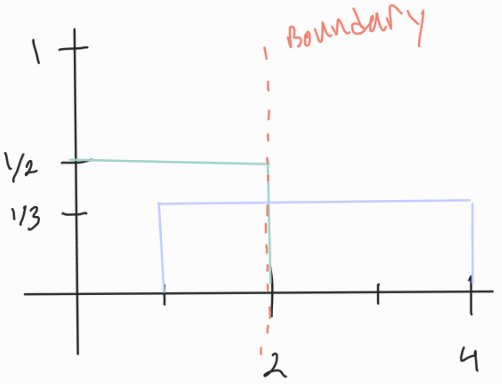
$$= \sqrt{1/2 \cdot 1/2} \cdot e^{-(0.1499)}$$

$$1/2 \cdot 0.8608 = 0.4304$$

Problem 4

Two one-dimensional distributions are uniform in [0,2] for ω_1 and [1,4] for ω_2, and $P_1=P_2=0.5$.

1. Find the Bayes boundary for minimum error, and compute the probability of error.
2. Find the Bayes decision boundary if $\lambda_{11}=\lambda_{22}=0$ and $\lambda_{12}=2\lambda_{21}$.



$$\text{Area} = L \cdot h = 1$$

$$2 \cdot \tfrac{1}{2} = 1$$

$$3 \cdot \tfrac{1}{3} = 1$$

$$P(\text{error}) = \int_{-\infty}^{\infty} P(\text{error}|x)\, p(x)\, dx$$

$$P(\text{error}) = \int_a^b P(\omega_1|x)\, P(x)\, dx$$

$$\frac{P(x|\omega_1)}{P(x|\omega_2)} > \frac{P(\omega_2)}{P(\omega_1)}$$

$$> 1$$

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

$$= \int_1^2 p(x|\omega)\, P(\omega)\, dx =$$

$$\int_1^2 \tfrac{1}{2} \cdot \tfrac{1}{2}\, dx = \tfrac{1}{4}2 - \tfrac{1}{4} = \tfrac{1}{4}$$

$$\int_1^2 \tfrac{1}{3} \cdot \tfrac{1}{2}\, dx = \tfrac{1}{6}2 - \tfrac{1}{6} = \tfrac{1}{6} \leftarrow \text{Min} = P(\text{error})$$

Decide ω_1 *if* $\dfrac{P(\mathbf{x}\,|\,\omega_1)}{P(\mathbf{x}\,|\,\omega_2)} > \dfrac{\lambda_{12}-\lambda_{22}}{\lambda_{21}-\lambda_{11}}\cdot\dfrac{P(\omega_2)}{P(\omega_1)}$ otherwise decide ω_2

$$\frac{P(x|\omega_1)}{P(x|\omega_2)} > \frac{2\cancel{\lambda_{21}} - 0}{\cancel{\lambda_{21}} - 0} \cdot \cancel{\frac{(1/2)}{(1/2)}}$$

$$\frac{P(x|\omega_1)}{P(x|\omega_2)} > 2$$

$$P(x|\omega_1) > 2\,P(\omega|\omega_2) \quad \text{or} \quad \frac{P(x|\omega_1)}{2} > P(x|\omega_2)$$

Problem 5

For a two-class recognition problem with salmon ($\omega = 1$) and sea bass ($\omega = 2$), suppose we have two features $\mathbf{x} = (x1, x2)$ and the two class-conditional densities, $p(x|\omega = 1)$ and $p(x|\ \omega = 2)$, are 2D Gaussian distributions centered at points (4, 16) and (16, 4) respectively with the same covariance matrix $\mathbf{\Sigma} = 4\mathbf{I}$ (with $\mathbf{I}$ is the identity matrix). Suppose the priors are $P(\omega = 1) = 0.6$ and $P(\omega = 2) = 0.4$.

1. Suppose we use a Bayes decision rule, write the two discriminant functions $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$.
2. Derive the equation for the decision boundary $g_1(\mathbf{x}) = g_2(\mathbf{x})$. Draw the boundary on the feature space (the 2D plane).

$$g_i(\mathbf{x}) = {w_i}^T\mathbf{x} + W_{i0}$$

$$w_i = \frac{1}{\sigma^2}\boldsymbol{\mu}_i$$

$$W_{i0} = -\frac{1}{2\sigma^2}{\boldsymbol{\mu}_i}^T\boldsymbol{\mu}_i + lnP(\omega_i)$$

$$W_1 = \frac{1}{4}\begin{pmatrix}4\\16\end{pmatrix} = \begin{pmatrix}1\\4\end{pmatrix}$$

$$W_{10} = -\frac{1}{2(4)}\begin{pmatrix}4 & 16\end{pmatrix}\begin{pmatrix}4\\16\end{pmatrix} + \ln(0.6)$$

$$= -34.5108$$

$$g_1(x) = X_1 + 4X_2 - 34.51$$

$$W_2 = \frac{1}{4}\begin{pmatrix}16\\4\end{pmatrix} = \begin{pmatrix}4\\1\end{pmatrix}$$

$$W_{20} = -\frac{1}{2(4)}\begin{pmatrix}16 & 4\end{pmatrix}\begin{pmatrix}16\\4\end{pmatrix} + \ln(0.6)$$

$$= -34.9163$$

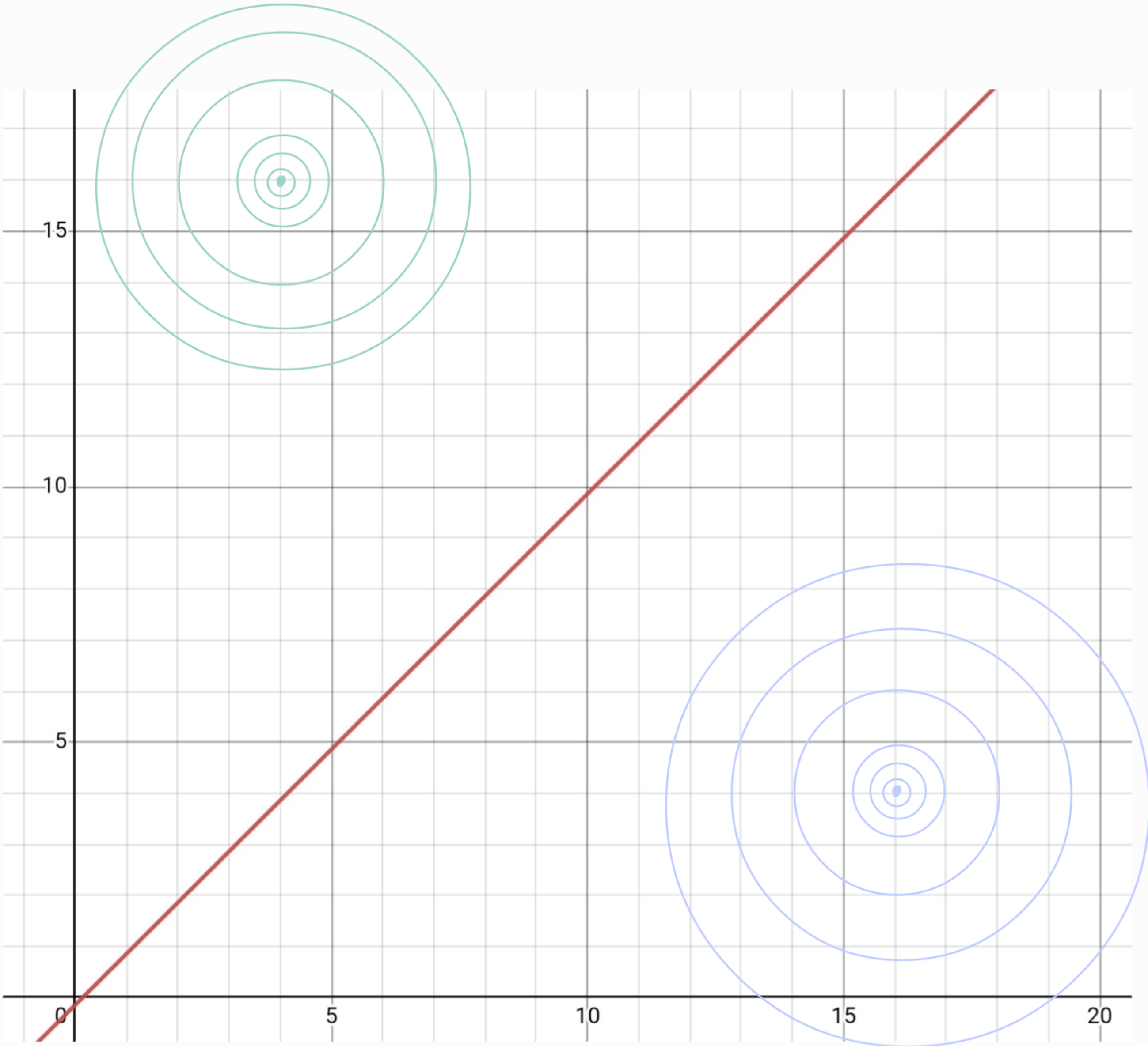
$$g_2(x) = 4X_1 + X_2 - 34.92$$

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

$$0 = g_1(x) - g_2(x)$$

$$= -3X_1 + 3X_2 + 0.4055$$

$$X_2 = \frac{3X_1 - 0.4055}{3} = X_1 - 0.1352$$


15
10
5
0
5
10
15
20