演習問題（等式制約の最適化）

4. 次のような5つのデータが与えられたとする.

$$(-3,-2),\ (1,-1),\ (0,0),\ (3,2),\ (-1,1)$$

1) 平均値ベクトル，および共分散行列 $\mathbf{C}$ を求めよ.

2) $\mathbf{C}$ の最大固有値と，その固有値に対する固有ベクトルを求めよ.

3) データの散布図を描け．また，この散布図に重ねて2)で求めた固有ベクトルを描け.

1) 平均ベクトルおよび共分散行列は以下のように書ける.

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{C} = \frac{1}{5}\begin{bmatrix} 20 & 10 \\ 10 & 10 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$$

2) 以下の固有値問題を解く.

$$\begin{aligned} & \mathbf{Cw} = \lambda \mathbf{w} \\ & \Leftrightarrow \det(\mathbf{C} - \lambda \mathbf{I}) = 0 \\ & \Leftrightarrow (4 - \lambda)(2 - \lambda) - 2 \cdot 2 = 0 \\ & \Leftrightarrow \lambda^2 - 6\lambda + 4 = 0 \\ & \therefore \lambda = 3 \pm \sqrt{5} \end{aligned}$$

このとき，最大の固有値として $\lambda = 3 + \sqrt{5}$ を得る.

最大固有値に対する固有ベクトル $\mathbf{w} = (x, y)^{\mathrm{T}}$ は以下の式を満たす.

$$(\mathbf{C} - \lambda\mathbf{I})\mathbf{w} = 0$$

$$\Leftrightarrow \begin{bmatrix} 4 - 5.24 & 2 \\ 2 & 2 - 5.24 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} -1.24x + 2y = 0 \\ 2x - 3.24y = 0 \end{cases}$$

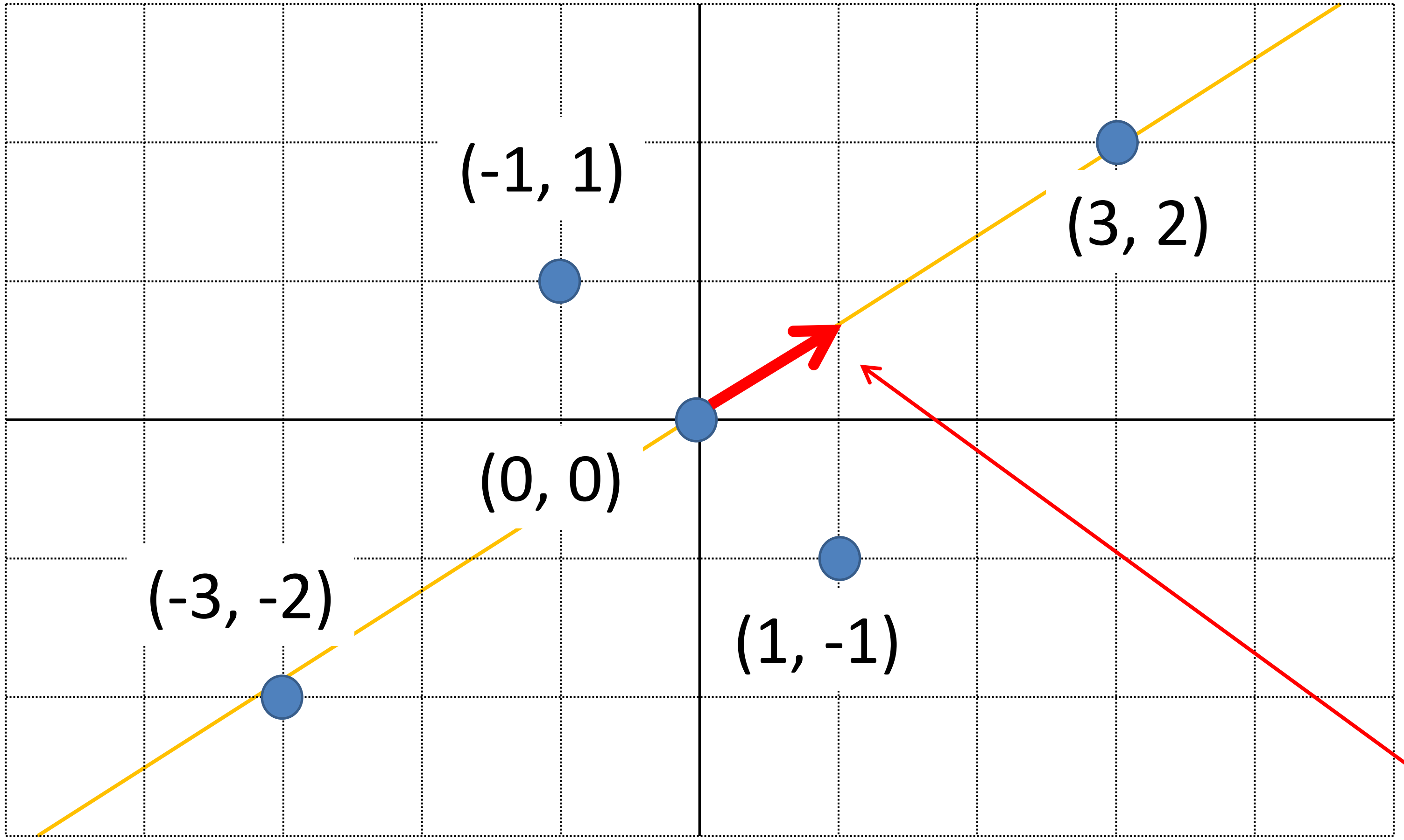
ここで, $\sqrt{5} = 2.24$ とした.

このとき, 求める固有ベクトル $\mathbf{w}$ は以下のように得られる.

$$\mathbf{w} = t\begin{bmatrix} 1 \\ 0.62 \end{bmatrix}$$

ここで, t は0ではない任意の実数である.

3) 全サンプルおよび固有ベクトルは以下のように図示できる.



$$\mathbf{w} = \begin{bmatrix} 1 \\ 0.62 \end{bmatrix}$$

演習問題（等式制約の最適化）

5．クラスαのデータとして，4つのデータ，

$$\boldsymbol{x}_1^\alpha=\begin{pmatrix}-5\\0\end{pmatrix},\ \boldsymbol{x}_2^\alpha=\begin{pmatrix}1\\4\end{pmatrix},\ \boldsymbol{x}_3^\alpha=\begin{pmatrix}-2\\1\end{pmatrix},\ \boldsymbol{x}_4^\alpha=\begin{pmatrix}-2\\3\end{pmatrix}$$

が与えられ，クラスβのデータとして，4つのデータ，

$$\boldsymbol{x}_1^\beta=\begin{pmatrix}2\\-1\end{pmatrix},\ \boldsymbol{x}_2^\beta=\begin{pmatrix}2\\-3\end{pmatrix},\ \boldsymbol{x}_3^\beta=\begin{pmatrix}5\\0\end{pmatrix},\ \boldsymbol{x}_4^\beta=\begin{pmatrix}-1\\-4\end{pmatrix}$$

が与えられている.

1) クラス毎に，平均ベクトル $\boldsymbol{\mu}_\alpha$， $\boldsymbol{\mu}_\beta$， 共分散行列 $\mathbf{C}_\alpha$, $\mathbf{C}_\beta$ を求めよ.

2) 級内分散行列 $\mathbf{C}_W$ を， ２つのクラスの共分散行列の平均として定義するとき， $\mathbf{C}_W$ を求めよ.

3) 級間分散行列 $\mathbf{C}_{\mathrm{B}}$ を，それぞれのクラスの平均ベクトルの共分散行列として定義するとき，$\mathbf{C}_{\mathrm{B}}$ を求めよ.

4) ベクトル $\mathbf{w} = (w_1\ w_2)^{\mathrm{T}}$ を用いて，$z_i = \mathbf{x}_i{}^{\mathrm{T}}\mathbf{w}$ と座標変換するとき，変換後の座標空間における級間分散 σ_{B}^2 と級内分散 σ_{W}^2 を，$\mathbf{C}_{\mathrm{B}}$，$\mathbf{C}_{\mathrm{W}}$，$\mathbf{w}$ の式で表せ.

5) 級間分散 σ_{B}^2 と級内分散 $\sigma_W{}^2$ の比 $\sigma_{\mathrm{B}}^2/\sigma_{\mathrm{W}}^2$ を最大化するように $\mathbf{w}$ を定めるとき，変換後の座標空間において，同じクラスのデータが纏まり，違うクラスのデータが離れる分布を作るため，座標空間は識別に適したものとなる．このような空間を作る，$\mathbf{C}_{\mathrm{B}}$，$\mathbf{C}_{\mathrm{W}}$，$\mathbf{w}$ の関係を導け.

6) 5) の条件を満たす $\mathbf{w}$ を実際に求めよ.

7) データの散布図を描け．この散布図に重ねて $\mathbf{w}$ を描け.

1) 各クラスの平均ベクトル，共分散行列は以下のように得られる.

$$\boldsymbol{\mu}_\alpha = \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \mathbf{C}_\alpha = \frac{1}{4}\begin{bmatrix} 18 & 12 \\ 12 & 10 \end{bmatrix} = \begin{bmatrix} 4.5 & 3 \\ 3 & 2.5 \end{bmatrix}$$

$$\boldsymbol{\mu}_\beta = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \mathbf{C}_\beta = \frac{1}{4}\begin{bmatrix} 18 & 12 \\ 12 & 10 \end{bmatrix} = \begin{bmatrix} 4.5 & 3 \\ 3 & 2.5 \end{bmatrix}$$

2) 級内分散行列は各クラスの共分散行列の平均として

$$\mathbf{C}_\mathrm{W} = \frac{1}{2}(\mathbf{C}_\alpha + \mathbf{C}_\beta) = \begin{bmatrix} 4.5 & 3 \\ 3 & 2.5 \end{bmatrix}$$

3) 級間分散行列は各クラスの平均ベクトルの共分散行列として

$$\mathbf{C}_\mathrm{B} = \frac{1}{2}\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}^T - \begin{bmatrix} 0 \\ 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -4 \\ -4 & 4 \end{bmatrix}$$

4) 変換後の部分空間における級内分散と級間分散は

$$\sigma_{\mathrm{W}}^2 = \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathrm{W}}\mathbf{w}$$

$$\sigma_{\mathrm{B}}^2 = \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathrm{B}}\mathbf{w}$$

のように書ける.

5) 変換後の部分空間において，級内分散一定という条件のもと，級間分散を最大にする問題を解けばよい．つまり，

$$E = \mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathrm{B}}\mathbf{w} - \lambda\left(\mathbf{w}^{\mathrm{T}}\mathbf{C}_{\mathrm{W}}\mathbf{w} - const.\right)$$

を$\mathbf{w}$で微分して0とおくと，以下の関係が得られる.

$$\mathbf{C}_{\mathrm{B}}\mathbf{w} - \lambda\mathbf{C}_{\mathrm{W}}\mathbf{w} = \mathbf{0} \Leftrightarrow \left(\mathbf{C}_{\mathrm{W}}^{-1}\mathbf{C}_{\mathrm{B}} - \lambda\mathbf{I}\right)\mathbf{w} = \mathbf{0}$$

6) 以下の固有値問題を解く．

$$\left(\mathbf{C}_\mathrm{W}^{-1}\mathbf{C}_\mathrm{B}-\lambda\mathbf{I}\right)\mathbf{w}=\mathbf{0}\Leftrightarrow \det\left(\mathbf{C}_\mathrm{W}^{-1}\mathbf{C}_\mathrm{B}-\lambda\mathbf{I}\right)=0$$
$$\Leftrightarrow (9.8-\lambda)(13.3-\lambda)-(-9.8)\cdot(-13.3)=0$$
$$\Leftrightarrow \lambda(\lambda-23.1)=0$$

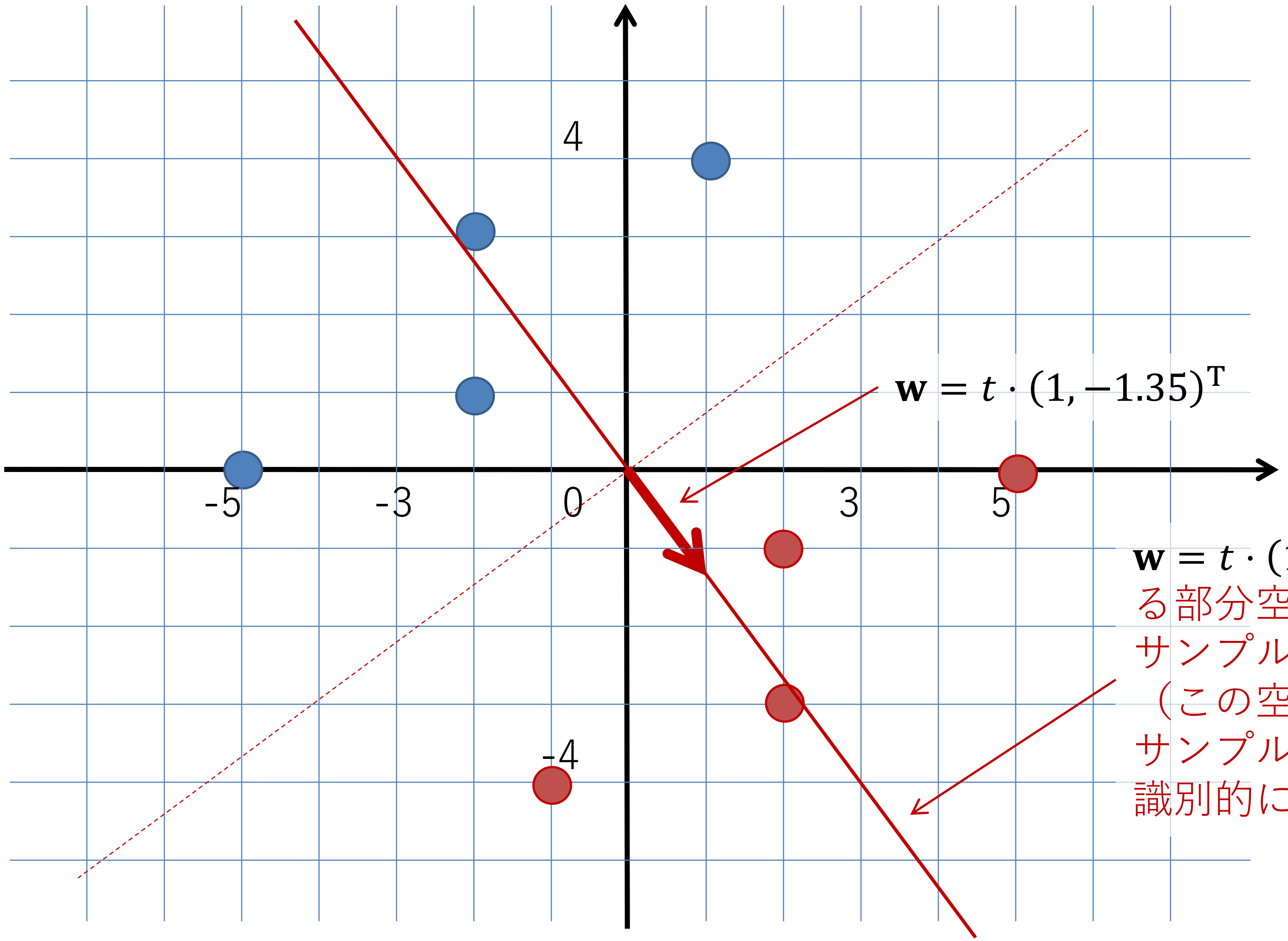
このとき，固有値として $\lambda=23.1$ を得る．

したがって，ベクトル $\mathbf{w}=(x,y)^\mathrm{T}$ は以下の式を満たす．

$$\left(\mathbf{C}_\mathrm{W}^{-1}\mathbf{C}_\mathrm{B}-\lambda\mathbf{I}\right)\mathbf{w}=0$$
$$\Leftrightarrow \begin{bmatrix}9.8-23.1 & -9.8\\ -13.3 & 13.3-23.1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}\Leftrightarrow \begin{cases}-13.3x-9.8y=0\\ -13.3x-9.8y=0\end{cases}$$

これを解けば，$\mathbf{w}=t\cdot(1,\ -1.35)^T$ が得られる（tは非零の実数）．

7)



$\mathbf{w} = t \cdot (1,\ -1.35)^{\mathrm{T}}$ が張る部分空間は各クラスのサンプルを識別的にする.（この空間に射影されたサンプルはクラスごとに識別的になる）