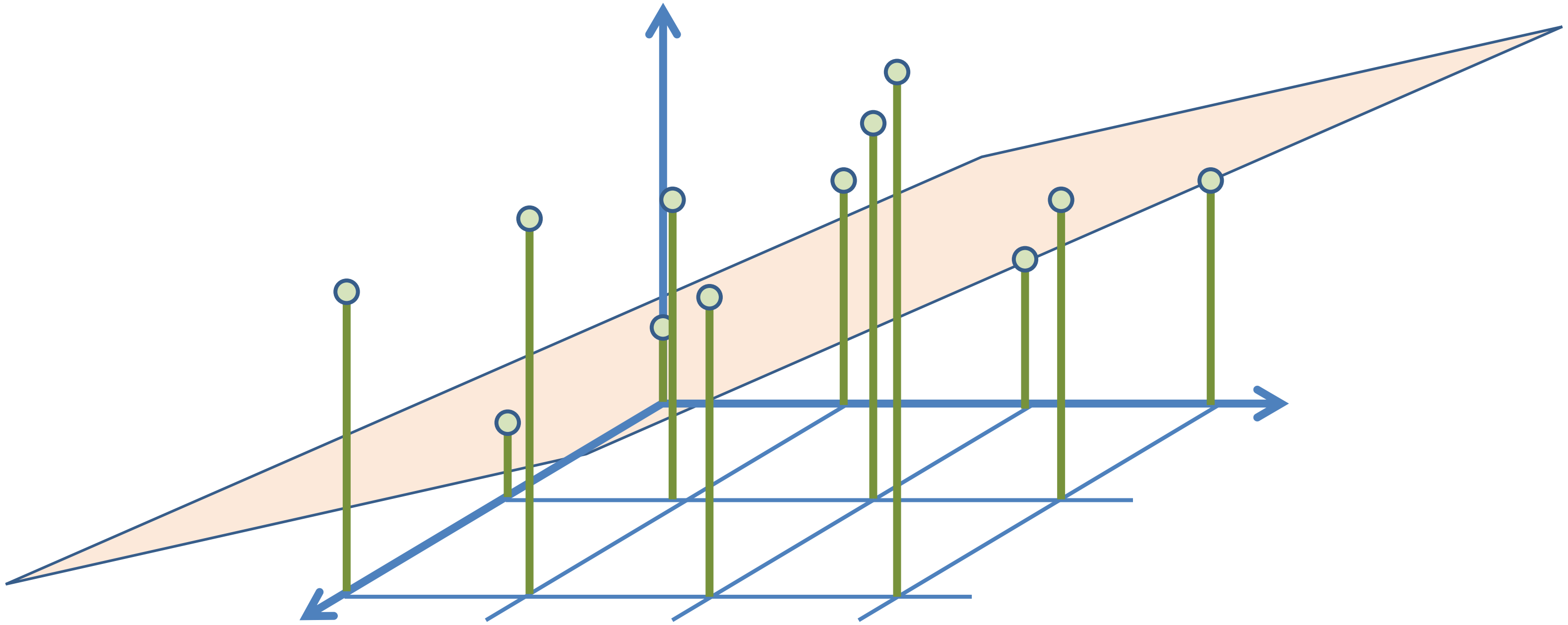
演習問題と解答

§ 演習問題 (線形回帰と最小二乗法)

1. 次の12点から，線形回帰モデルを，最小二乗基準で求めよ。

$$(0,0,1),(0,1,3),(0,2,2),(0,3,3)$$
$$(1,0,1),(1,1,4),(1,2,5),(1,3,4)$$
$$(2,0,4),(2,1,5),(2,2,4),(2,3,7)$$



解答 1.

Matlabのプログラム例

```
x = [0,0,0,0,1,1,1,1,2,2,2,2;0,1,2,3,0,1,2,3,0,1,2,3];
y = [1,3,2,3,1,4,5,4,4,5,4,7];
nd = 12;

X = [x;ones(1,nd)];

a = inv(X*X')*X*y'
---------------
a =
   1.37500
   0.76667
   1.05833
```

§ 演習問題（ロジスティック回帰）

2層のニューラルネット（パーセプトロン）を以下のように表現する。

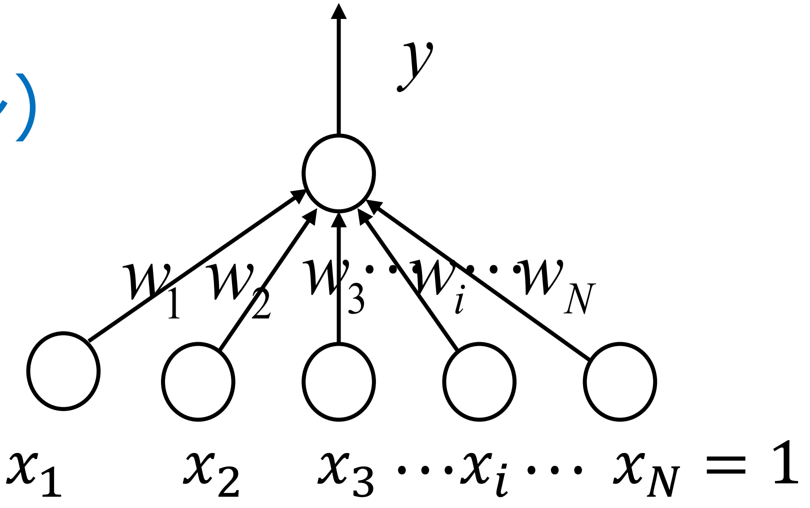
$\boldsymbol{x}^{(k)} = (x_1^{(k)}\ x_2^{(k)} \cdots x_i^{(k)} \cdots x_N^{(k)})^{\mathrm{T}}$
:第k入力ベクトル

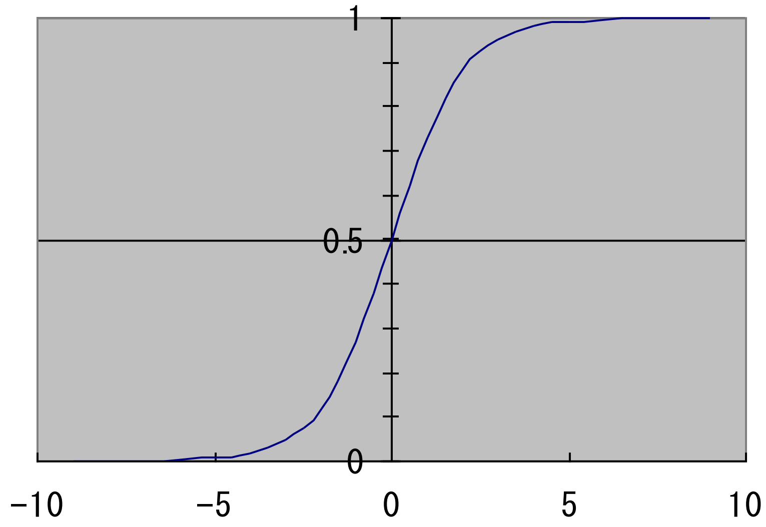
$y^{(k)}$:第k入力に対する出力

$$y^{(k)} = f\left(\sum_{i=1}^{N} w_i x_i^{(k)}\right)$$

$$f(z) = \frac{1}{1+\exp(-z)}$$

（カテゴリAであれば１，そうでなければ，0）





$\boldsymbol{w} = (w_1\ w_2 \cdots w_i \cdots w_N)^{\mathrm{T}}$

: 重みベクトル（パーセプトロンのパラメータ）

$t^{(k)}$: 第k入力に対する正解カテゴリ（カテゴリAであれば 1 そうでなければ 0）

M : データの個数

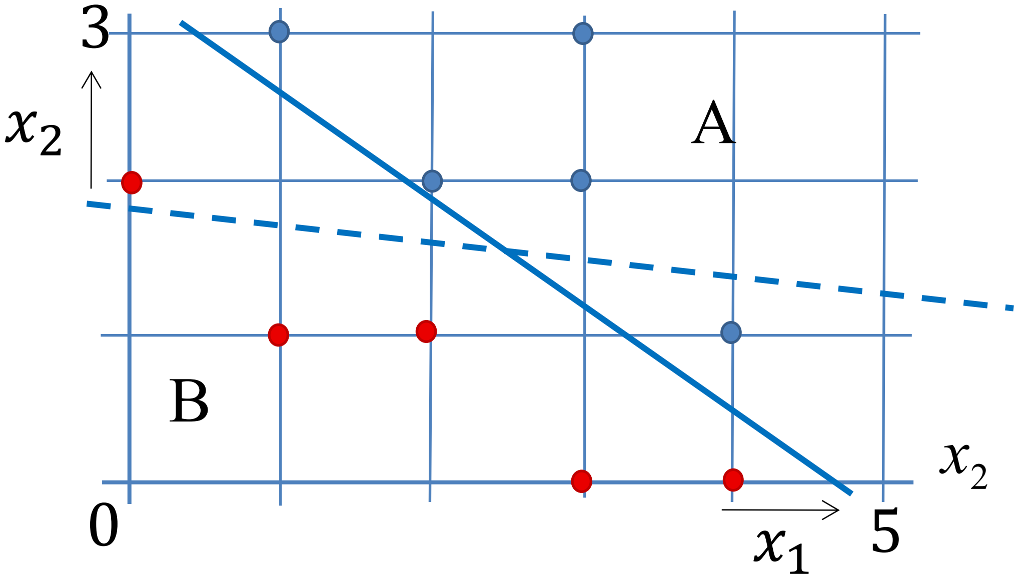
1) 目的関数を

$$H(\boldsymbol{w}) = \sum_{j=1}^{M} \frac{1}{2} \left(y^{(j)} - t^{(j)} \right)^2$$

$$= \sum_{j=1}^{M} \frac{1}{2} \left(f\left(\sum_{i=1}^{N} w_i x_i^{(j)} \right) - t^{(j)} \right)^2, \quad f(z) = \frac{1}{1+e^{-z}}$$

として定義するとき， $\boldsymbol{w}$ を最急降下法を用いて求めるアルゴリズムを書け。

2)学習データとして以下のものが与えられるとき， $\boldsymbol{w}$ を求めよ。
データの分布図に重ねて，wが決めるカテゴリの境界線を描け。



$x_3 = 1$

$$X = \left(\boldsymbol{x}^{(k)}\right)^T = \begin{pmatrix} 1 & 2 & 3 & 3 & 4 & 0 & 1 & 2 & 3 & 4 \\ 3 & 2 & 2 & 3 & 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^T$$

$$T = \left(t^{(k)}\right)^T = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T$$

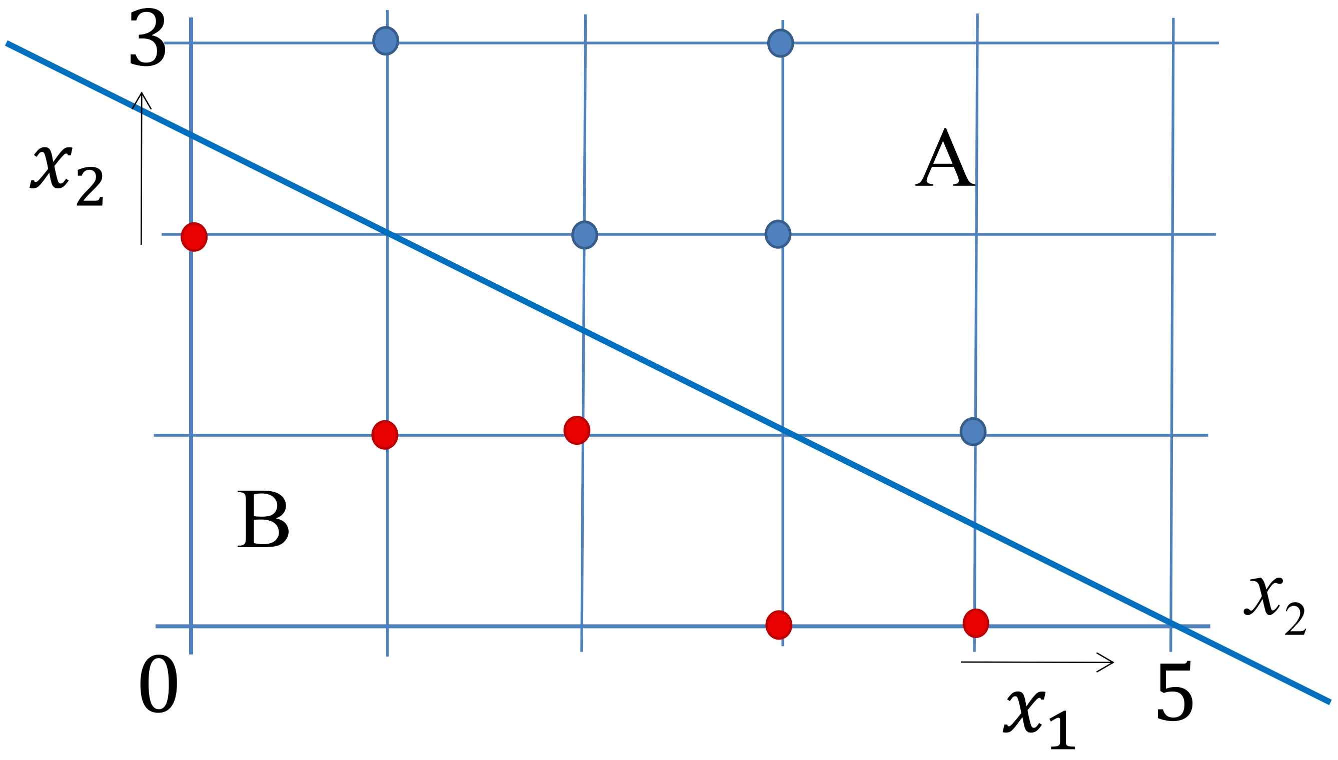
解答

```
function main
  alpha = 0.2;
  TH = 1.0e-7;
  NRPT = 100000;
  nd =10;
  data =
[1,2,3,3,4,0,1,2,3,4;3,2,2,3,1,2,1,1,0,0]';
  true = [1,1,1,1,1,0,0,0,0,0]';
  D = [data,ones(nd,1)];
  w = [0,0,0]';

  pErr = Inf;
  for i=1:NRPT
    estm = sigm(D*w);
    err = estm - true;
    G = err.*estm.*(1-estm).*D;
    for j=1:nd
      w = w - alpha*G(j,:)';
    end
    cErr = sqrt(err'*err)/nd;
    if (abs(cErr-pErr) < TH)
      break
    end
    pErr = cErr;
  end
  w
end

function a = sigm(x)
  a = 1./(1+exp(-x));
end
```

$$w = \begin{pmatrix} 3.7 \\ 7.4 \\ -18.5 \end{pmatrix}$$



$x_3 = 1$

§ 演習問題（いろいろな勾配法）

1. 関数

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T \begin{pmatrix} 6 & 1 \\ 1 & 2 \end{pmatrix} \boldsymbol{x} - \begin{pmatrix} 2 \\ 1 \end{pmatrix}^T \boldsymbol{x}$$

の　最小値を，
共役勾配法，ニュートン法のそれぞれで求めよ。

解答 1.

共役勾配法（Matlabでの実装例）

```
A=[6,1;1,2];
b=[2,1]';
x = [0,0]';
r = -A*x+b;
p = r;
for i=1:10
    c = (p'*b)/(p'*A*p);
    x = x + c*p;
    r = -A*x+b;
    if abs(r)<=0.0001
        break
    end
    alp = -(p'*A*r)/(p'*A*p);
    p = r + alp*p;
end
x
```

```
x =
   0.27273
   0.36364
```

$$f(x) = \frac{1}{2}x^T\begin{pmatrix}6 & 1\\1 & 2\end{pmatrix}x - \begin{pmatrix}2 & 1\end{pmatrix}x$$

$$\nabla f(x) = \begin{pmatrix}6 & 1\\1 & 2\end{pmatrix}x - \begin{pmatrix}2\\1\end{pmatrix}$$

$$x_0 = \begin{pmatrix}0\\0\end{pmatrix}$$

$$r_0 = -\begin{pmatrix}6 & 1\\1 & 2\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix} + \begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}$$

$$p_1 = r_0 = \begin{pmatrix}2\\1\end{pmatrix}$$

$$c_1 = \frac{p_1' b}{p_1' A p_1} = \frac{\begin{pmatrix}2 & 1\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}}{\begin{pmatrix}2 & 1\end{pmatrix}\begin{pmatrix}6 & 1\\1 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}} = \frac{1}{6}$$

$$x_1 = x_0 + c_1 p_1 = \frac{1}{6}\begin{pmatrix}2\\1\end{pmatrix}$$

$$r_1 = -Ax_1 + b = -\frac{1}{6}\begin{pmatrix}6 & 1\\1 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix} + \begin{pmatrix}2\\1\end{pmatrix} = \frac{1}{6}\begin{pmatrix}-1\\2\end{pmatrix}$$

$$\alpha_1 = -\frac{{p_1}^T A r_1}{{p_1}^T A p_1} = -\frac{1}{6}\frac{\begin{pmatrix}2 & 1\end{pmatrix}\begin{pmatrix}6 & 1\\1 & 2\end{pmatrix}\begin{pmatrix}-1\\2\end{pmatrix}}{\begin{pmatrix}2 & 1\end{pmatrix}\begin{pmatrix}6 & 1\\1 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}} = \frac{1}{36}$$

$$p_2 = r_1 + \alpha_1 p_1 = \frac{1}{6}\begin{pmatrix}-1\\2\end{pmatrix} + \frac{1}{36}\begin{pmatrix}2\\1\end{pmatrix} = \frac{1}{36}\begin{pmatrix}-4\\13\end{pmatrix}$$

$$c_2 = \frac{p_2' b}{p_2' A p_2} = \frac{1}{36}\frac{\begin{pmatrix}-4 & 13\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}}{\begin{pmatrix}-4 & 13\end{pmatrix}\begin{pmatrix}6 & 1\\1 & 2\end{pmatrix}\begin{pmatrix}-4\\13\end{pmatrix}} = \frac{6}{11}$$

$$x_2 = x_1 + c_2 p_2 = \frac{1}{11}\begin{pmatrix}3\\4\end{pmatrix}$$

ニュートン法（Matlabでの実装例，二次系なので繰り返しにならない）

```
A=[6,1;1,2];
b=[2,1]';
x = [0,0]';
df = A*x-b;
H = A;
x = x - inv(H)*df;
x
```

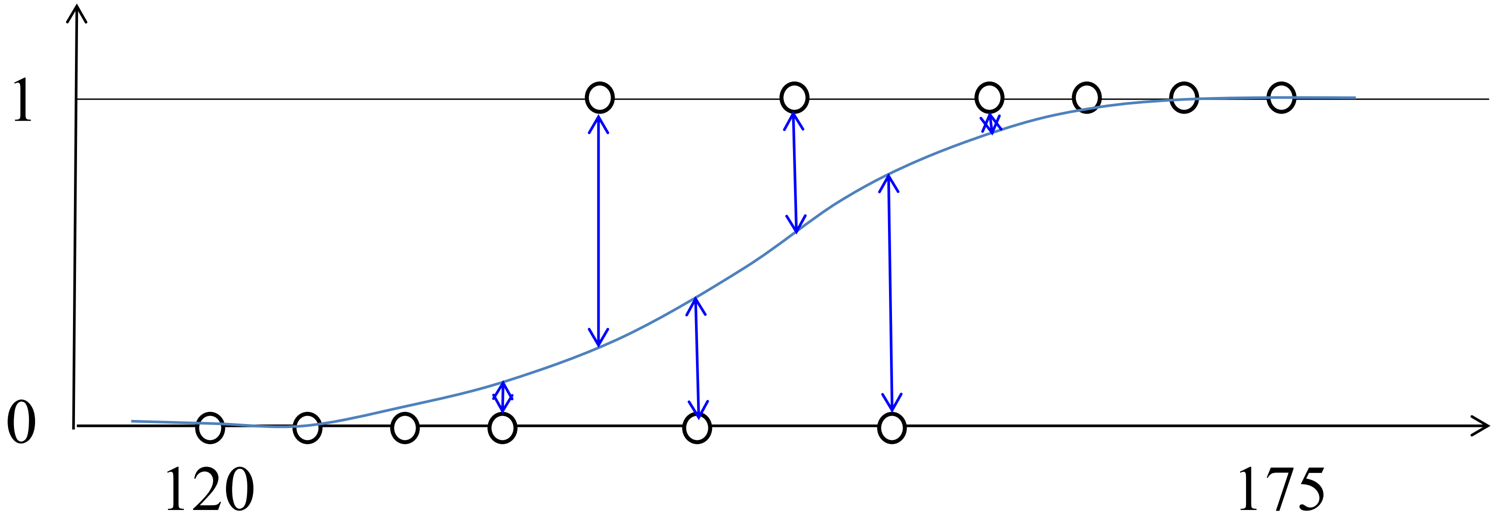
```
x =
   0.27273
   0.36364
```

演習問題（いろいろな勾配法）

2. 下記の学習データに対し，ロジスティック関数（シグモイド関数）を当てはめることを考える。

（1.20，0）（1.25，0）（1.30，0）（1.35，0）（1.40，1）（1.45，0）
（1.50，1）（1.55，0）（1.60，1）（1.65，1）（1.70，1）（1.75，1）

(1)二乗誤差最小化するパラメタを，勾配法を用いて求めよ。
(2)二乗誤差最小化するパラメタを，ニュートン法を用いて求めよ。



解答 2.

(1) 例題のとおり

(2)
```
function testNewton
  alp = 0.2;
  TH = 1.0e-10;
  NRPT = 100000;
  nd =12;
  data = 1.2:0.05:1.75;
  true =
[0,0,0,0,1,0,1,0,1,1,1,1]';
  D = [data;ones(1,nd)]';
  w = [0,0]';

  pErr = Inf;
```

```
for i=1:NRPT
  estm = sigm(D*w);
  err = estm - true;
  G = err.*estm.*(1-estm).*D;
  dw = (ones(1,nd)*G)'/nd;
  c = (-3*estm.*estm+2*estm.*(1+true)-true).*estm.*(1-estm);
  H = (c.*D)'*D/(ones(1,nd)*c);
  w = w - inv(H) * dw;
  cErr = sqrt(err'*err)/nd;
  if (abs(cErr-pErr) < TH)
    break
  end
  pErr = cErr;
 end
 w
end
```

```
function a = sigm(x)
  a = 1./(1+exp(-x));
end
```

演習問題（いろいろな勾配法）

3. 勾配法における，近似解の更新方法（更新ベクトルの定め方,学習率の定め方 ）には様々な方法がある。

   授業で触れなかったものについて，どのような方法があるか各自調査せよ。

   また，そのいくつかを実装し，先のロジスティック回帰の問題（演習問題2）に適用して，収束の様子を比べよ。

§ 演習問題（等式制約の最適化）

1.

条件

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - 4\begin{pmatrix}0 & 1\end{pmatrix}\boldsymbol{x} + 7 = 0 \qquad (1)$$

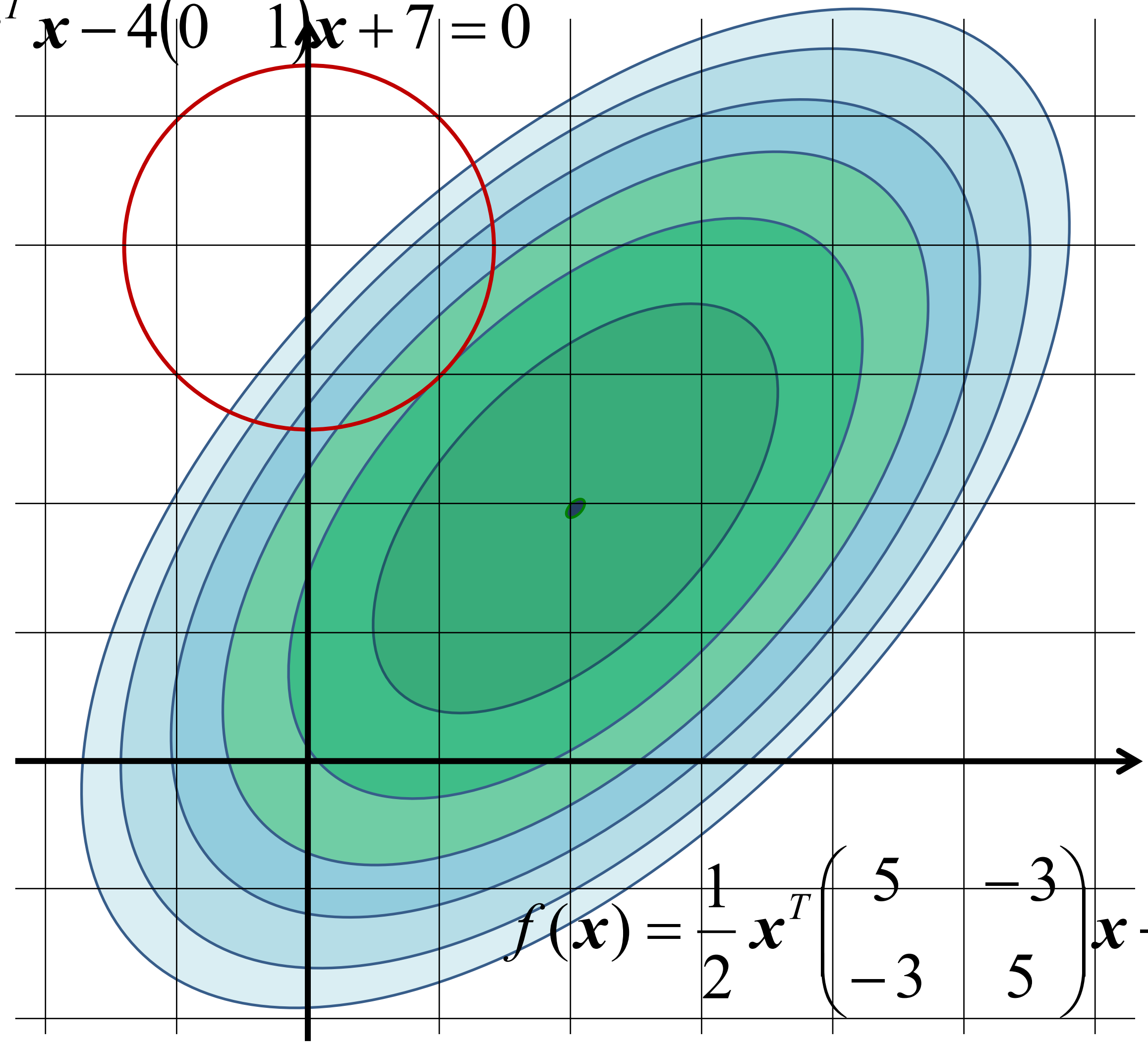
の下で，

関数，

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix}5 & -3\\ -3 & 5\end{pmatrix}\boldsymbol{x} - 4\begin{pmatrix}1\\ 1\end{pmatrix}^T\boldsymbol{x} \qquad (2)$$

の最小値を与えるxを求めよ。

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - 4\begin{pmatrix}0 & 1\end{pmatrix}\boldsymbol{x} + 7 = 0$$



$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix}5 & -3\\ -3 & 5\end{pmatrix}\boldsymbol{x} - 4\begin{pmatrix}1\\ 1\end{pmatrix}^T\boldsymbol{x} = c$$

解答：

評価関数を次のように置く

$$L(\boldsymbol{x},\lambda)=f(\boldsymbol{x})+\lambda g(\boldsymbol{x})$$

$$=\frac{1}{2}\boldsymbol{x}^T\begin{pmatrix}5 & -3\\ -3 & 4\end{pmatrix}\boldsymbol{x}-4\begin{pmatrix}1\\1\end{pmatrix}^T\boldsymbol{x}+\lambda\left(\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}-4\begin{pmatrix}0 & 1\end{pmatrix}\boldsymbol{x}+7\right) \qquad (3)$$

xとλで偏微分して0と置く

$$\frac{\partial}{\partial \boldsymbol{x}}L(\boldsymbol{x},\lambda)=f(\boldsymbol{x})+\lambda g(\boldsymbol{x})=\begin{pmatrix}5+\lambda & -3\\ -3 & 5+\lambda\end{pmatrix}\boldsymbol{x}-4\begin{pmatrix}1\\ 1+\lambda\end{pmatrix}=0 \qquad (4)$$

$$\frac{\partial}{\partial \lambda}L(\boldsymbol{x},\lambda)=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}-4\begin{pmatrix}0 & 1\end{pmatrix}\boldsymbol{x}+7=0 \qquad (5)$$

上の連立方程式を解く。

(4)より，

$$\boldsymbol{x} = 4\begin{pmatrix} 5+\lambda & -3 \\ -3 & 5+\lambda \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ 1+\lambda \end{pmatrix}$$

$$= \frac{4}{(5+\lambda)(5+\lambda)-9}\begin{pmatrix} 5+\lambda & 3 \\ 3 & 5+\lambda \end{pmatrix}\begin{pmatrix} 1 \\ 1+\lambda \end{pmatrix} = \frac{4}{(\lambda+8)}\begin{pmatrix} 4 \\ \lambda+4 \end{pmatrix} \qquad (6)$$

$$= \frac{4}{(\eta+2)}\begin{pmatrix} 1 \\ \eta+1 \end{pmatrix}$$

ただし，

$$4\eta = \lambda$$

(6)を(5)に代入する。

$$\frac{4^2}{2(\eta+2)^2}\begin{pmatrix}1 & \eta+1\end{pmatrix}\begin{pmatrix}1\\ \eta+1\end{pmatrix}-\frac{4^2}{(\eta+2)}\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1\\ \eta+1\end{pmatrix}+7=0 \qquad (7)$$

η について整理して，これを解くと，

$$\begin{aligned}&8\left((\eta+1)^2+1\right)-16(\eta+2)(\eta-1)+7(\eta+2)^2\\ &=\left(8\eta^2+16\eta+16\right)-\left(16\eta^2+48\eta+32\right)+\left(7\eta^2+28\eta+28\right)\\ &=-\eta^2-4\eta+12=-(\eta+6)(\eta-2)=0\end{aligned}$$

$$\eta=2,-6 \qquad (8)$$

(8)を(6)に代入して，

$$\boldsymbol{x} = \frac{4}{(\eta+2)}\begin{pmatrix} 1 \\ \eta+1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \end{pmatrix}$$

これを(2)に代入して，

$$f\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 3 \end{pmatrix}\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} - 4\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$
$$= 0$$

$$f\begin{pmatrix} -1 \\ 5 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 & 5 \end{pmatrix}\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} -1 \\ 5 \end{pmatrix} - 4\begin{pmatrix} 1 & 1 \end{pmatrix}\begin{pmatrix} -1 \\ 5 \end{pmatrix}$$
$$= 64$$

求める$\boldsymbol{x}$は，$(1,3)^T$ であり，最小値は，0　となる。

（正のηに対応するものが最小値を与える。）

演習問題（等式制約の最適化）

2.

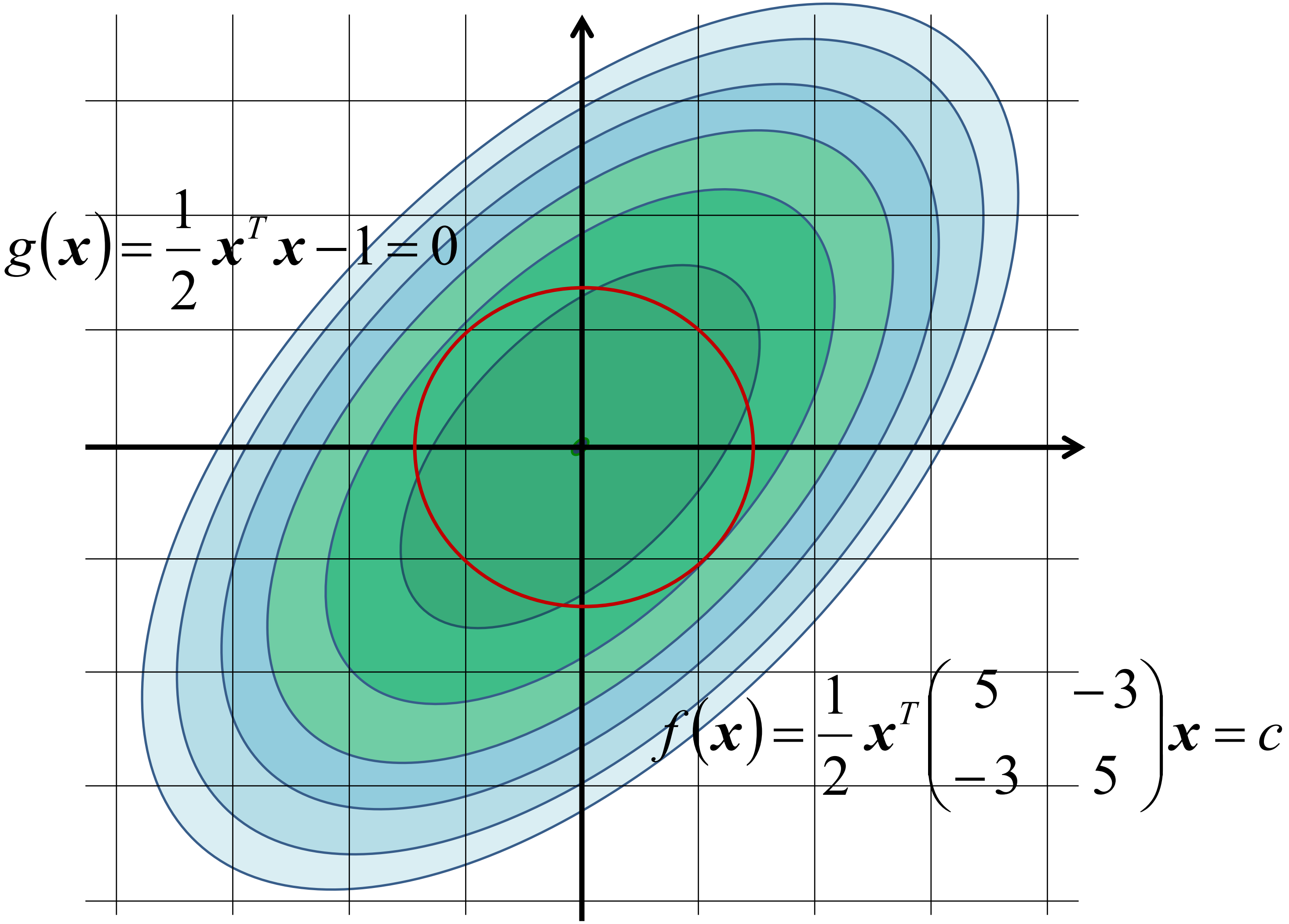
条件,

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - 1 = 0 \qquad (1)$$

のもとで，関数

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\boldsymbol{x} \qquad (2)$$

の最小値を与えるxを求めよ。


$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - 1 = 0$
$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\boldsymbol{x} = c$

解答：

評価関数を次のように置く

$$L(\boldsymbol{x},\lambda)=f(\boldsymbol{x})-\lambda g(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T\begin{pmatrix}5 & -3\\ -3 & 5\end{pmatrix}\boldsymbol{x}-\lambda\left(\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}-1\right) \qquad (3)$$

$\boldsymbol{x}$とλで偏微分して0と置く

$$\frac{\partial}{\partial \boldsymbol{x}}L(\boldsymbol{x},\lambda)=\begin{pmatrix}5-\lambda & -3\\ -3 & 5-\lambda\end{pmatrix}\boldsymbol{x}=0 \qquad (4)$$

$$\frac{\partial}{\partial \lambda}L(\boldsymbol{x},\lambda)=-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}+1=0 \qquad (5)$$

上の連立方程式を解く。

(4)より，

$$\det\begin{pmatrix} 5-\lambda & -3 \\ -3 & 5-\lambda \end{pmatrix} = 0$$

$$\lambda^2 - 10\lambda + 16 = (\lambda - 8)(\lambda - 2) = 0$$

$$\lambda = 2,8$$

(4) より，

$$\begin{pmatrix} 5-\lambda & -3 \end{pmatrix}\boldsymbol{x} = 0$$

$\lambda = 2$ のとき，

$$\begin{pmatrix} 3 & -3 \end{pmatrix}\boldsymbol{x} = 0$$

$$\boldsymbol{x} = k\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\lambda = 8$ のとき，

$$\begin{pmatrix} -3 & -3 \end{pmatrix}\boldsymbol{x} = 0$$

$$\boldsymbol{x} = k\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

(1)より，

$$\frac{k^2}{2}\begin{pmatrix}1 & \pm 1\end{pmatrix}\begin{pmatrix}1\\ \pm 1\end{pmatrix}-1=0$$

$$k=\pm 1$$

よって，

$$\boldsymbol{x}=\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}$$

このとき，$f(\boldsymbol{x})$は，

$$f\begin{pmatrix}1\\1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1 & 1\end{pmatrix}\begin{pmatrix}5 & -3\\-3 & 5\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=2$$

$$f\begin{pmatrix}1\\-1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1 & -1\end{pmatrix}\begin{pmatrix}5 & -3\\-3 & 5\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix}=32$$

よって，求める　$\boldsymbol{x}$　は，$(1,1)^{\mathrm{T}}$　最小値は，2

演習問題（等式制約の最適化）

3. 4つのデータ，

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

が与えられている。

1) このデータの共分散行列 C を求めよ。

2) ベクトル $\boldsymbol{w} = (w_1\ w_2)^T$を用いて，$z_i = {\boldsymbol{x}_i}^T\boldsymbol{w}$ と座標変換するとき，$\|\boldsymbol{w}\| = 1$ の条件の下で，z_iの分散を最も大きくするように, $\boldsymbol{w}$ を定めることにするとき，Cと $\boldsymbol{w}$ の関係を導け。

3) 2)の条件を満たす $\boldsymbol{w}$ を実際に求めよ。

4) データの散布図を描け。この散布図に重ねて$\boldsymbol{w}$ を描け。

解答：

1)

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

$$N = 4, \quad \boldsymbol{\mu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$X = \begin{pmatrix} 5 & -3 & 3 & -5 \\ 0 & -4 & 4 & 0 \end{pmatrix}^T$$

$$C = \frac{1}{4}\begin{pmatrix} 5 & -3 & 3 & -5 \\ 0 & -4 & 4 & 0 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ -3 & -4 \\ 3 & 4 \\ -5 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 17 & 6 \\ 6 & 8 \end{pmatrix}$$

2)

$$L(\boldsymbol{w}, \lambda) = \boldsymbol{w}^T C \boldsymbol{w} - \lambda(\boldsymbol{w}^T \boldsymbol{w} - 1)$$

$$\frac{\partial}{\partial \boldsymbol{w}} L(\boldsymbol{w}, \lambda) = (C - \lambda I)\boldsymbol{w}$$

3)

$$\det(C - \lambda I) = 0, \qquad \begin{vmatrix} 17 - \lambda & 6 \\ 6 & 8 - \lambda \end{vmatrix} = 0$$

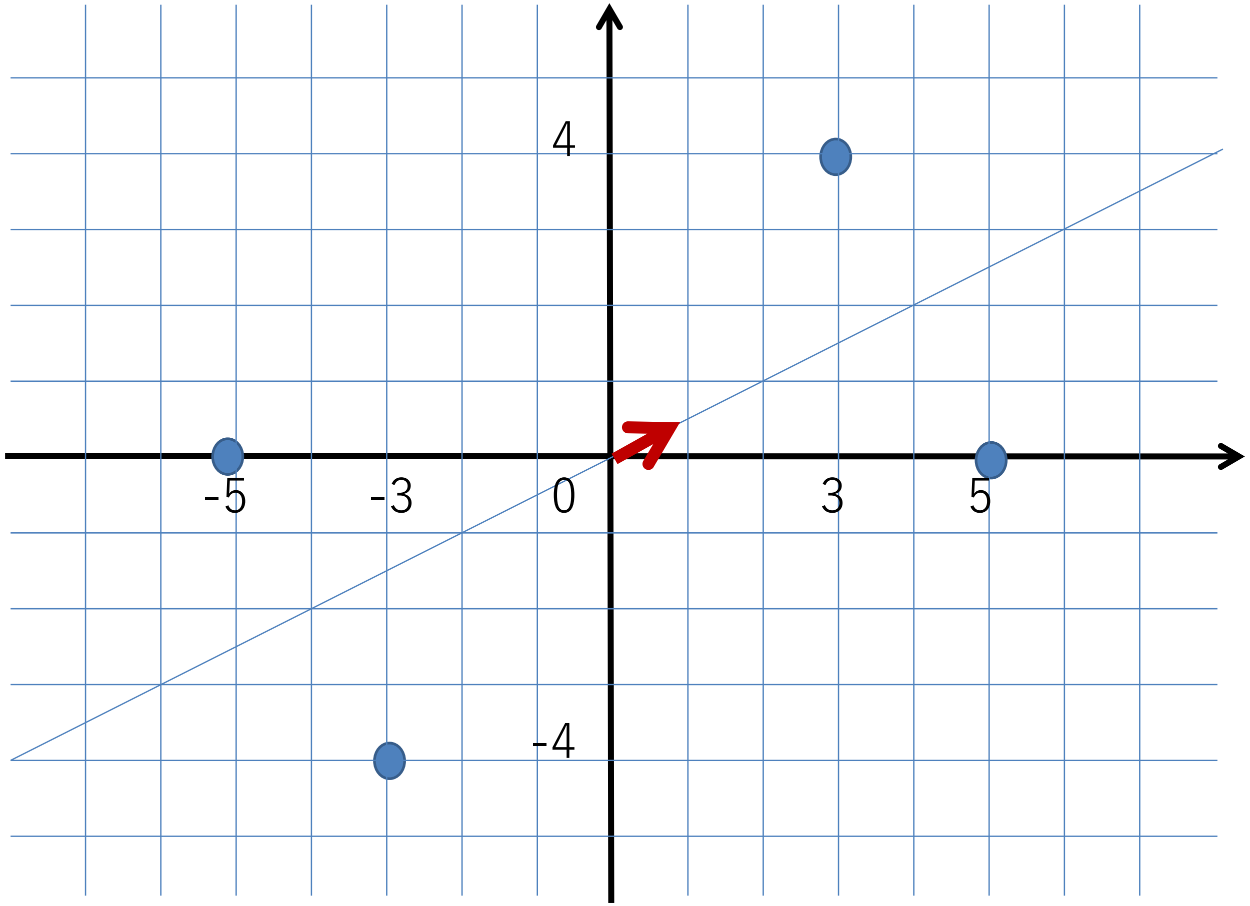
$$(17 - \lambda)(8 - \lambda) - 6 \cdot 6 = (\lambda - 20)(\lambda - 5) = 0$$

最大固有値は $\lambda = 20$

最大固有値に対応した固有ベクトルは，$(17 - 20 \quad 6) = (-1,2)$に直交する単位ベクトル

$$\boldsymbol{w} = \frac{1}{\sqrt{5}} L \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

4)



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　C　を求めよ。

2)　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2.2で求めた固有ベクトルを描け。

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$C_{\alpha}$, C_{β}を求めよ。

2) 級内分散行列 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, C_W を求めよ。

3) 級間分散行列 C_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， C_B を求めよ。

4) ベクトル $\boldsymbol{w} = (w_1\ w_2)^T$を用いて，$z_i = \boldsymbol{x}_i{}^T \boldsymbol{w}$ 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 $\sigma_B{}^2$ と級内分散 $\sigma_W{}^2$ を，C_B，C_W，$\boldsymbol{w}$ の式で表せ。

5) 級間分散 $\sigma_B{}^2$ と級内分散 $\sigma_W{}^2$ の比 $\sigma_B{}^2/\sigma_W{}^2$ を最大化するように $\boldsymbol{w}$ 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，C_B，C_W，$\boldsymbol{w}$ の関係を導け。

6) 5)の条件を満たす$\boldsymbol{w}$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$\boldsymbol{w}$を描け。

演習問題（不等式制約の最適化）

1.

条件

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - \begin{pmatrix} 2 & 2 \end{pmatrix}\boldsymbol{x} + 2 \leq 0 \qquad (1)$$

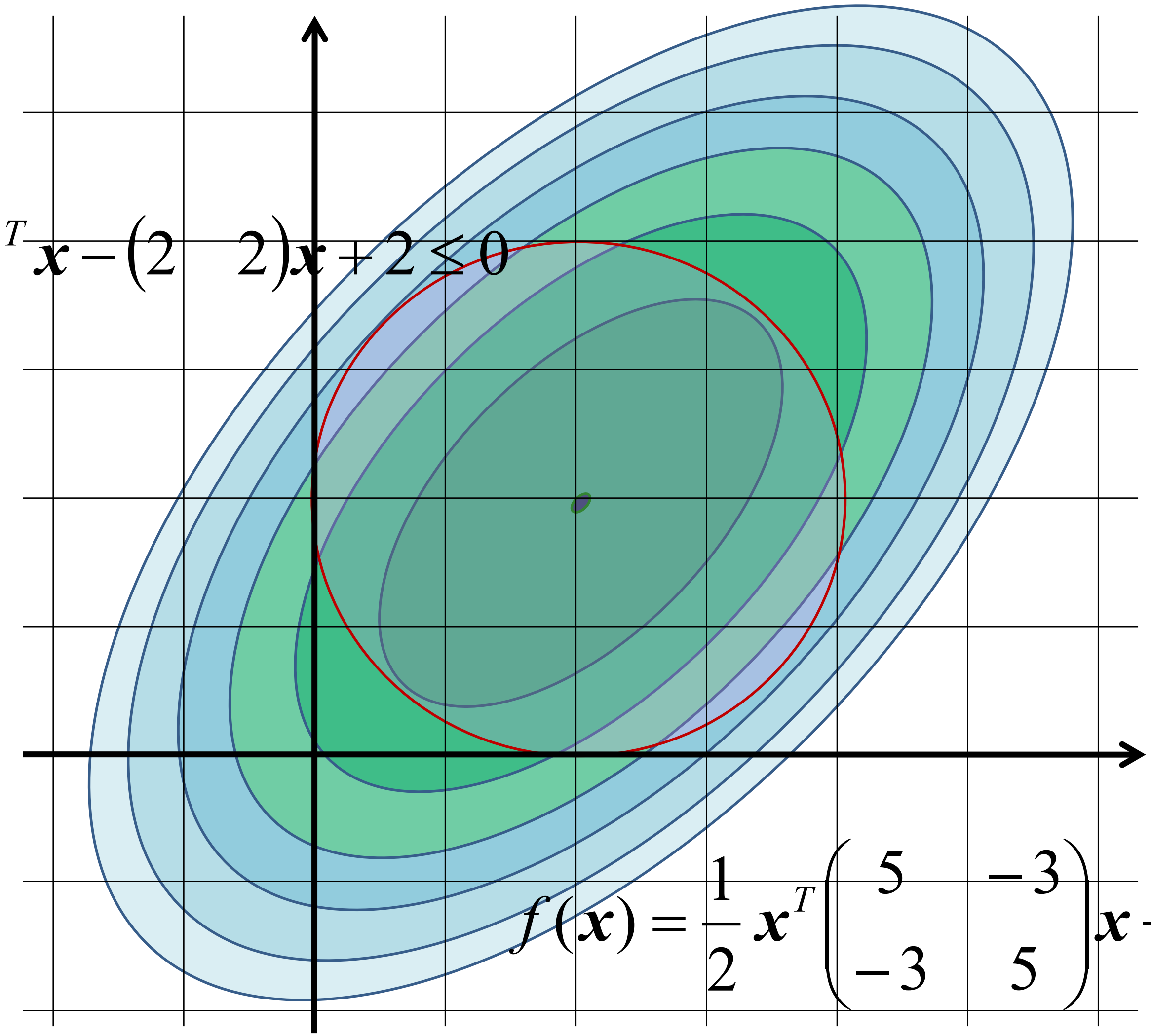
の下で，

関数，

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\boldsymbol{x} - 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}^T\boldsymbol{x} \qquad (2)$$

の最小値を与える*x*を求めよ。

$$g(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{x} - \begin{pmatrix} 2 & 2 \end{pmatrix}\boldsymbol{x} + 2 \le 0$$



$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix}\boldsymbol{x} - 4\begin{pmatrix} 1 \\ 1 \end{pmatrix}^T\boldsymbol{x} = c$$

演習問題（SVM）

1. SVMの説明にある，式(14) を導け。

2. サポートベクタに対するラグランジュ係数を$\boldsymbol{\alpha}^0$，マージンをρとするとき，

$$\boldsymbol{\psi}^{0^T}\boldsymbol{\psi}^0 = \sum_{i=1}^{m} \alpha_i^0$$

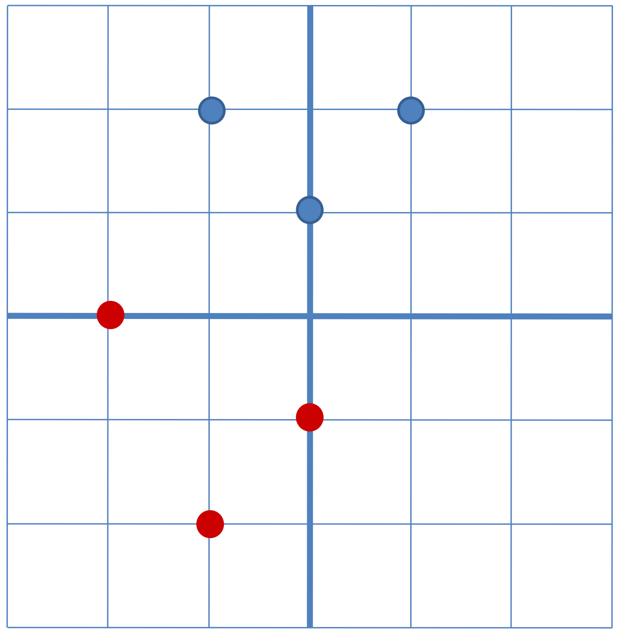
$$\rho(\boldsymbol{\psi}^0) = \frac{1}{\|\boldsymbol{\psi}^0\|}$$

$$L(\boldsymbol{\alpha}) < L(\boldsymbol{\alpha}^0) = \frac{1}{2}\frac{1}{\rho(\boldsymbol{\psi}^0)^2}, \quad \boldsymbol{\alpha} \neq \boldsymbol{\alpha}^0$$

を証明せよ。

演習問題（SVM）

3. カテゴリAとして，(1,2),(0,1),(-1,2)が，カテゴリBとして，(0,-1),(-1,-2),(-2,0) が与えられたとき，これらのカテゴリをマージン最大化基準で分離する直線を求めよ。



解答 3：

$\boldsymbol{x}_1 = (1,2,1)^T$, $\boldsymbol{x}_2 = (0,1,1)^T$, $\boldsymbol{x}_3 = (-1,2,1)^T$,
$\boldsymbol{x}_4 = (0,-1,1)^T$, $\boldsymbol{x}_5 = (-1,-2,1)^T$, $\boldsymbol{x}_6 = (-2,0,1)^T$,

$\boldsymbol{y} = (1,1,1,-1,-1,-1)^T$

$$R = (r_{ij}) = ((\boldsymbol{x}_i * \boldsymbol{x}_j)) = \begin{pmatrix} 6 & 3 & 4 & -1 & -4 & -1 \\ 3 & 2 & 3 & 0 & -1 & 1 \\ 4 & 3 & 6 & -1 & -2 & 3 \\ -1 & 0 & -1 & 2 & 3 & 1 \\ -4 & -1 & -2 & 3 & 6 & 3 \\ -1 & 1 & 3 & 1 & 3 & 5 \end{pmatrix}$$

$$Y = (y_i y_j) = \begin{pmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \end{pmatrix}$$

$$A = \left(y_i y_j r_{ij}\right) = R \circ Y$$

$$= \begin{pmatrix} 6 & 3 & 4 & -1 & -4 & -1 \\ 3 & 2 & 3 & 0 & -1 & 1 \\ 4 & 3 & 6 & -1 & -2 & 3 \\ -1 & 0 & -1 & 2 & 3 & 1 \\ -4 & -1 & -2 & 3 & 6 & 3 \\ -1 & 1 & 3 & 1 & 3 & 5 \end{pmatrix} \circ \begin{pmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 1 & 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 6 & 3 & 4 & 1 & 4 & 1 \\ 3 & 2 & 3 & 0 & 1 & -1 \\ 4 & 3 & 6 & 1 & 2 & -3 \\ 1 & 0 & 1 & 2 & 3 & 1 \\ 4 & 1 & 2 & 3 & 6 & 3 \\ 1 & -1 & -3 & 1 & 3 & 5 \end{pmatrix}$$

$$L(\boldsymbol{\alpha}) = -\frac{1}{2}\boldsymbol{\alpha}^T A\boldsymbol{\alpha} + \|\boldsymbol{\alpha}\|_1 \qquad \frac{\partial}{\partial\boldsymbol{\alpha}}L(\boldsymbol{\alpha}) = -A\boldsymbol{\alpha} + \mathbf{1}$$

$$\boldsymbol{\alpha}_{k+1} = \boldsymbol{\alpha}_k + c\frac{\partial}{\partial\boldsymbol{\alpha}}L(\boldsymbol{\alpha}_k)$$

$$\boldsymbol{\alpha}_0 = (0,0,0,0,0,0)^T$$

$$\frac{\partial}{\partial\alpha}L(\boldsymbol{\alpha}_\mathbf{0}) = -\begin{pmatrix} 6 & 3 & 4 & 1 & 4 & 1 \\ 3 & 2 & 3 & 0 & 1 & -1 \\ 4 & 3 & 6 & 1 & 2 & -3 \\ 1 & 0 & 1 & 2 & 3 & 1 \\ 4 & 1 & 2 & 3 & 6 & 3 \\ 1 & -1 & -3 & 1 & 3 & 5 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$c = 0.2$

$$\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_0 + 0.2 \frac{\partial}{\partial \boldsymbol{\alpha}} L(\boldsymbol{\alpha}_0) = (0.2, 0.2, 0.2, 0.2, 0.2, 0.2)^T$$

$$\frac{\partial}{\partial \boldsymbol{\alpha}} L(\boldsymbol{\alpha_1}) = -\begin{pmatrix} 6 & 3 & 4 & 1 & 4 & 1 \\ 3 & 2 & 3 & 0 & 1 & -1 \\ 4 & 3 & 6 & 1 & 2 & -3 \\ 1 & 0 & 1 & 2 & 3 & 1 \\ 4 & 1 & 2 & 3 & 6 & 3 \\ 1 & -1 & -3 & 1 & 3 & 5 \end{pmatrix} \begin{pmatrix} 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \\ 0.2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2.8 \\ -0.6 \\ -1.6 \\ -0.6 \\ -2.8 \\ -0.2 \end{pmatrix}$$

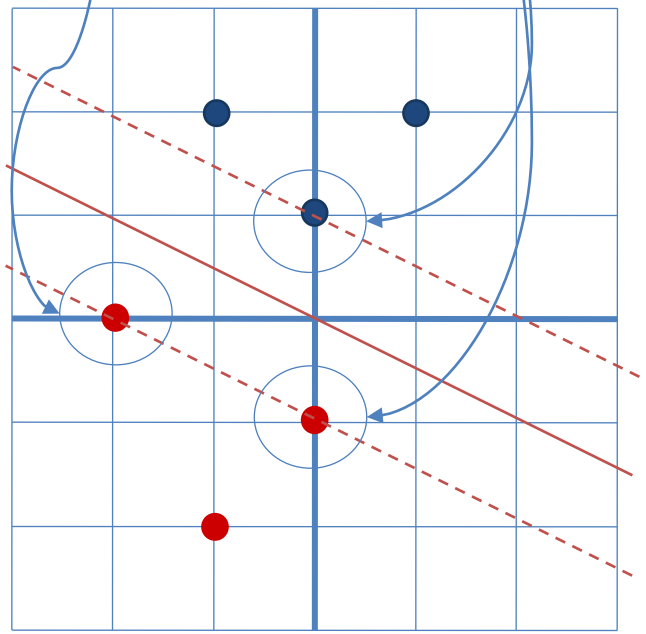
$$\boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_1 + 0.2 \frac{\partial}{\partial \boldsymbol{\alpha}} L(\boldsymbol{\alpha}_1) = (0, 0.08, 0, 0.08, 0, 0.16)^T$$

$$\frac{\partial}{\partial \boldsymbol{\alpha}} L(\boldsymbol{\alpha_2}) = -\begin{pmatrix} 6 & 3 & 4 & 1 & 4 & 1 \\ 3 & 2 & 3 & 0 & 1 & -1 \\ 4 & 3 & 6 & 1 & 2 & -3 \\ 1 & 0 & 1 & 2 & 3 & 1 \\ 4 & 1 & 2 & 3 & 6 & 3 \\ 1 & -1 & -3 & 1 & 3 & 5 \end{pmatrix} \begin{pmatrix} 0.00 \\ 0.08 \\ 0.00 \\ 0.08 \\ 0.00 \\ 0.16 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.52 \\ 1.00 \\ 1.16 \\ 0.68 \\ 0.20 \\ 0.20 \end{pmatrix}$$

$$\vdots$$

$$\boldsymbol{\alpha}_{19} = (0, 0.623, 0, 0.375, 0, 0.250)^T$$
$$\boldsymbol{\alpha}_{20} = (0, 0.624, 0, 0.375, 0, 0.250)^T$$
$$\boldsymbol{\alpha}_{21} = (0, 0.624, 0, 0.375, 0, 0.250)^T$$
$$\boldsymbol{\alpha}_{22} = (0, 0.625, 0, 0.375, 0, 0.250)^T$$

Support vectors



$$\boldsymbol{\psi}^0 = \sum_{i \in S} \alpha_i^{\ 0} y_i \boldsymbol{x}_i, \ \ S = \{2,4,6\}$$

$$\boldsymbol{\psi}^0 = 0.625 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - 0.375 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} - 0.250 \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 1.0 \\ 0 \end{pmatrix}$$

$$\boldsymbol{\psi}^0 \boldsymbol{x} = 0.5x + y = 0$$