

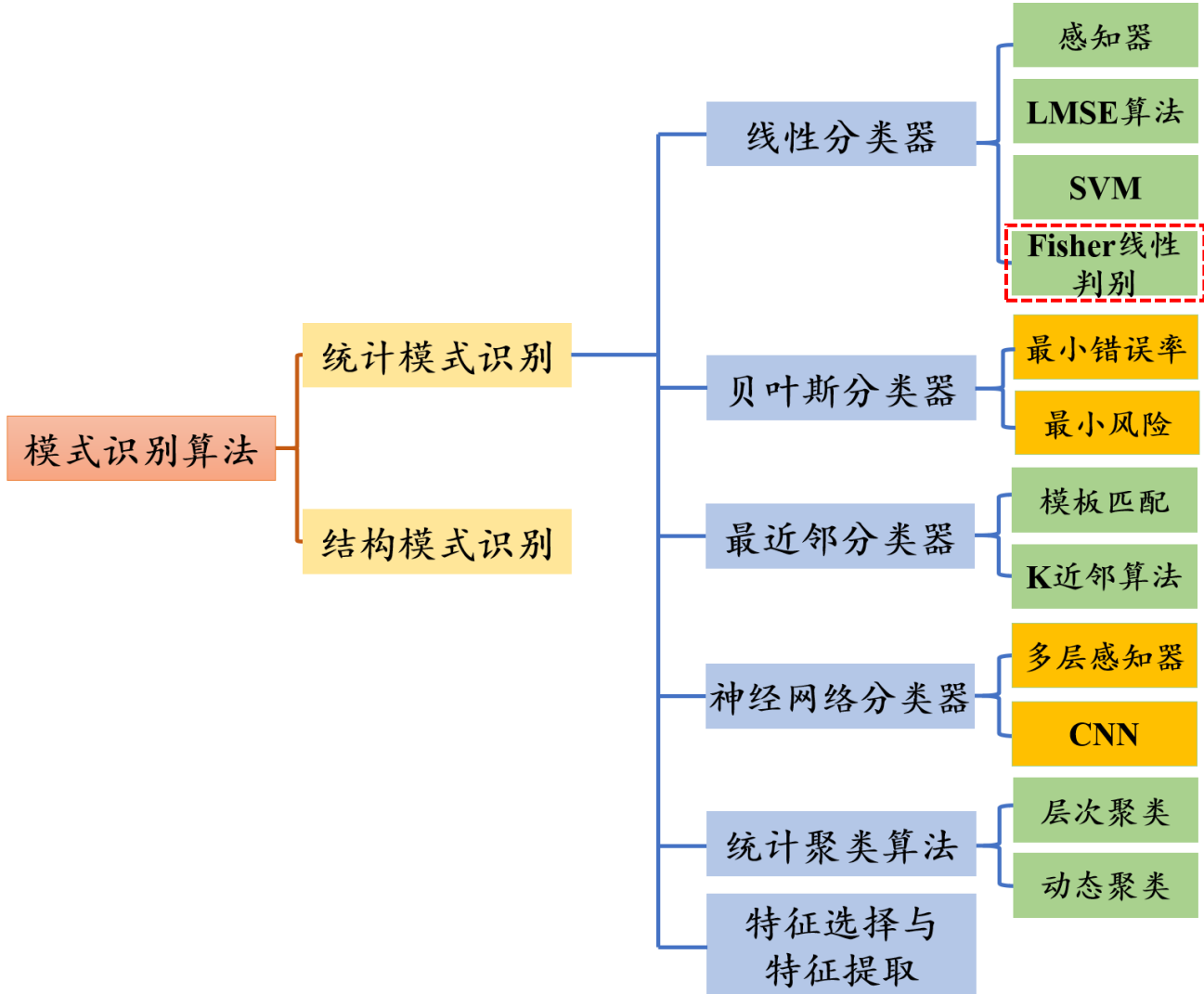
Fisher线性判别
张俊超

中南大学
航空航天学院





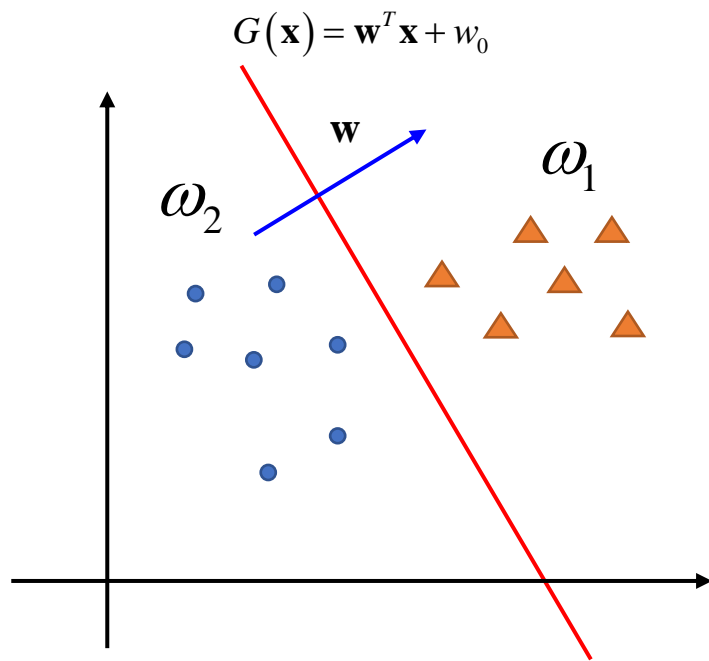
模式识别-线性分类器





线性分类器 – Fisher线性判别

Fisher线性判别：把样本**投影**到一个方向上，然后在这个**一维空间**确定分类的阈值。通过阈值点且与投影方向垂直的超平面就是分类面。

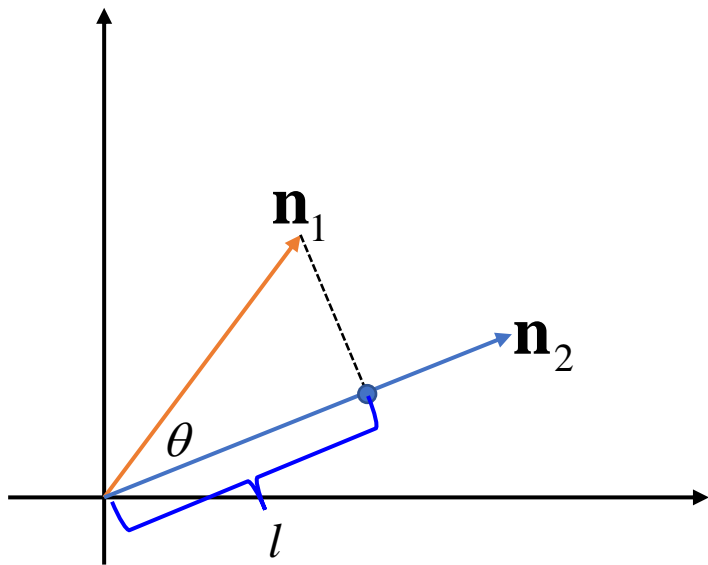


线性分类器 – Fisher线性判别

Q: 什么是投影?

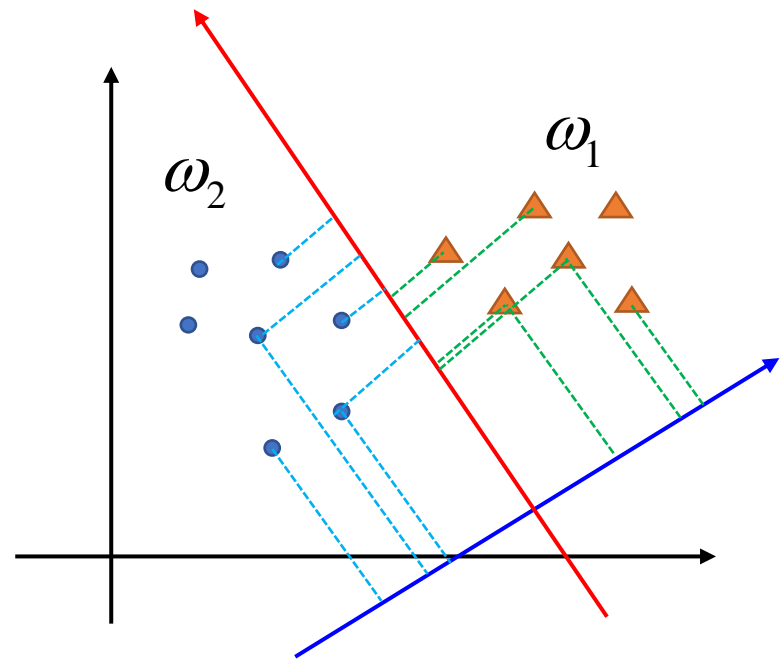
\mathbf{n}_1 在 \mathbf{n}_2 上的投影是 l

$$l = \|\mathbf{n}_1\| \cos \theta = \|\mathbf{n}_1\| \frac{\mathbf{n}_1^T \mathbf{n}_2}{\|\mathbf{n}_1\| \|\mathbf{n}_2\|} = \frac{\mathbf{n}_1^T \mathbf{n}_2}{\|\mathbf{n}_2\|}$$



Q: 如何选择投影方向?

Q: 哪一个投影方向最佳?



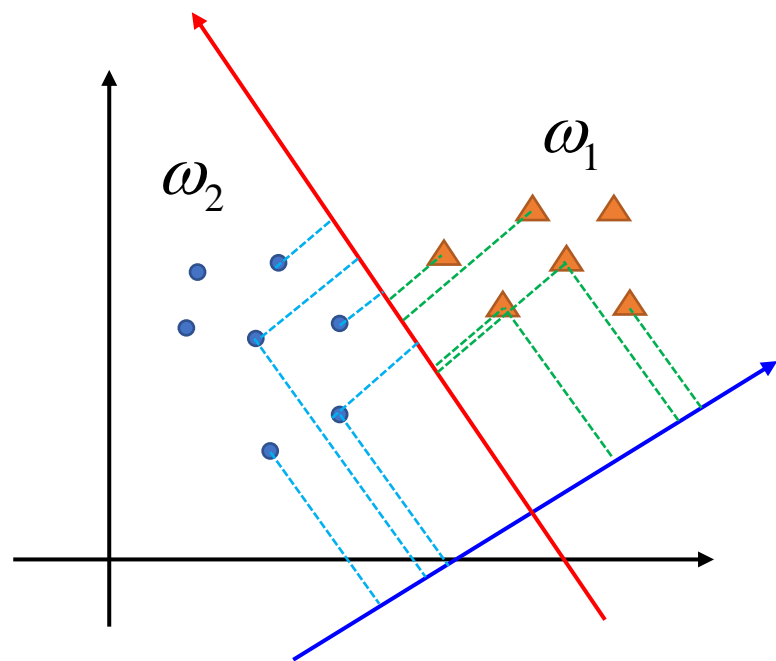


线性分类器 – Fisher线性判别

Fisher线性判别思想：

选择投影方向，使投影后两类相隔尽可能远，而同一类内部的样本又尽可能的聚集。

Q: 如何判断聚类程度





线性分类器 – Fisher线性判别

记号:

ω_1 的样本为 $\mathbb{X}_1 = \{\mathbf{x}_1^1, \mathbf{x}_2^1, \dots, \mathbf{x}_{N_1}^1\}$, ω_2 的样本为 $\mathbb{X}_2 = \{\mathbf{x}_1^2, \mathbf{x}_2^2, \dots, \mathbf{x}_{N_2}^2\}$

投影方向为: \mathbf{w} , 则投影以后的样本为: $y_i = \mathbf{w}^T \mathbf{x}_i$

定义:

原样本空间的【类内离散度矩阵】为:

$$\mathbf{S}_i = \sum_{\mathbf{x}_j^i \in \mathbb{X}_i} (\mathbf{x}_j^i - \mathbf{m}_i)(\mathbf{x}_j^i - \mathbf{m}_i)^T, \quad i = 1, 2$$

$$\mathbf{m}_i = \frac{1}{N_i} \sum_{\mathbf{x}_j^i \in \mathbb{X}_i} \mathbf{x}_j^i, \quad i = 1, 2$$

【总类内离散度矩阵】为: $\mathbf{S}_w = \mathbf{S}_1 + \mathbf{S}_2$ 对称矩阵

【类间离散度矩阵】为: $\mathbf{S}_b = (\mathbf{m}_1 - \mathbf{m}_2)(\mathbf{m}_1 - \mathbf{m}_2)^T$ 对称矩阵



线性分类器 – Fisher线性判别

投影后(一维空间)

$$\text{【均值】}\bar{m}_i = \frac{1}{N_i} \sum_{\mathbf{x}_j^i \in \mathbb{X}_i} \mathbf{w}^T \mathbf{x}_j^i = \mathbf{w}^T \mathbf{m}_i, \quad i = 1, 2$$

$$\text{【类内离散度】}\bar{s}_i^2 = \sum_{\mathbf{x}_j^i \in \mathbb{X}_i} \left(y_j^i - \bar{m}_i \right)^2, \quad i = 1, 2$$

$$\text{【总类内离散度】}\bar{s}_w = \bar{s}_1^2 + \bar{s}_2^2$$

$$\text{【类间离散度】}\bar{s}_b = \left(\bar{m}_1 - \bar{m}_2 \right)^2$$



*Fisher*线性判别准则:

$$\max J_F(\mathbf{w}) = \frac{\bar{s}_b}{\bar{s}_w} = \frac{\left(\bar{m}_1 - \bar{m}_2 \right)^2}{\bar{s}_1^2 + \bar{s}_2^2} = \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}$$



线性分类器 – Fisher线性判别

$$\max J_F(\mathbf{w}) = \frac{\overline{s_b}}{\overline{s_w}} = \frac{(\overline{m_1} - \overline{m_2})^2}{\overline{s_1^2} + \overline{s_2^2}} = \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}$$

\Downarrow

$$\max \{ \mathbf{w}^T \mathbf{S}_b \mathbf{w} \}, s.t. \mathbf{w}^T \mathbf{S}_w \mathbf{w} = c \neq 0$$

\Downarrow

$$L(\mathbf{w}, \lambda) = \mathbf{w}^T \mathbf{S}_b \mathbf{w} - \lambda (\mathbf{w}^T \mathbf{S}_w \mathbf{w} - c)$$

$$\Downarrow \frac{\partial L}{\partial \mathbf{w}} = 0$$

$$\mathbf{S}_b \mathbf{w}^* - \lambda \mathbf{S}_w \mathbf{w}^* = 0$$

$$\nabla_{\mathbf{x}} (\mathbf{x}^T \mathbf{A} \mathbf{x}) = (\mathbf{A} + \mathbf{A}^T) \mathbf{x}$$

若 \mathbf{S}_w 是非奇异的, 则 $\mathbf{S}_w^{-1} \mathbf{S}_b \mathbf{w}^* = \lambda \mathbf{w}^*$



线性分类器 – Fisher线性判别

$$\mathbf{S}_w^{-1} \mathbf{S}_b \mathbf{w}^* = \lambda \mathbf{w}^*$$

\Downarrow

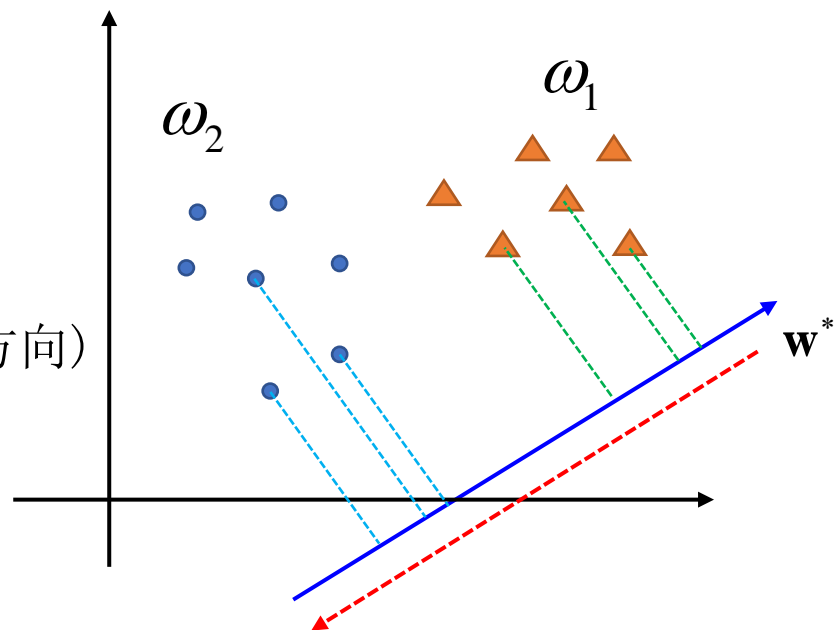
$$\lambda \mathbf{w}^* = \mathbf{S}_w^{-1} (\mathbf{m}_1 - \mathbf{m}_2) (\mathbf{m}_1 - \mathbf{m}_2)^T \mathbf{w}^*$$

\Downarrow

$$\lambda \mathbf{w}^* = \mathbf{S}_w^{-1} (\mathbf{m}_1 - \mathbf{m}_2) \cdot r(\text{标量, 不影响}\mathbf{w}^*\text{的方向})$$

\Downarrow

$$\mathbf{w}^* = \frac{r}{\lambda} \mathbf{S}_w^{-1} (\mathbf{m}_1 - \mathbf{m}_2)$$



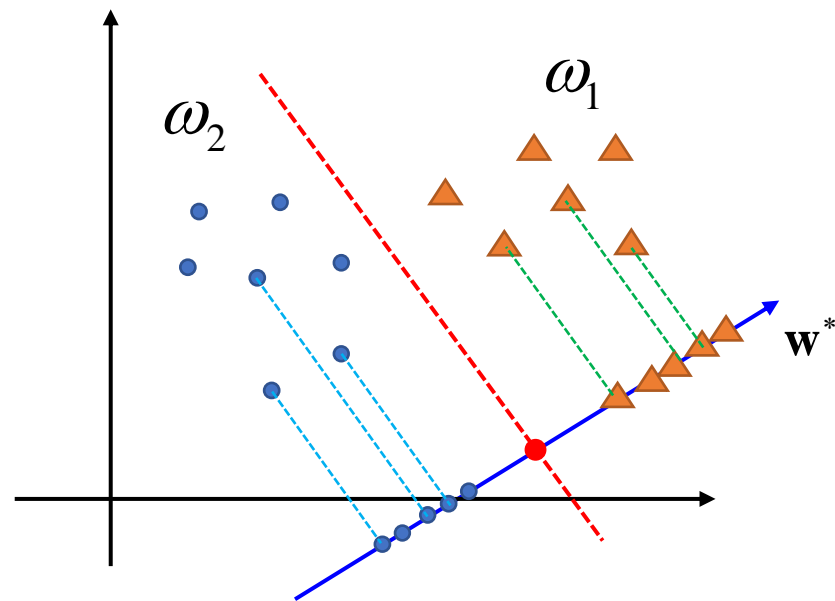
忽略常数的影响，常取 $\mathbf{w}^* = \mathbf{S}_w^{-1} (\mathbf{m}_1 - \mathbf{m}_2)$

Q: Fisher判别给出了最优投影方向，唯一吗？

Q: Fisher判别给出了最优投影方向，怎么求分类面呢？



线性分类器 – Fisher线性判别



分类准则：判断位于 d 的哪一侧

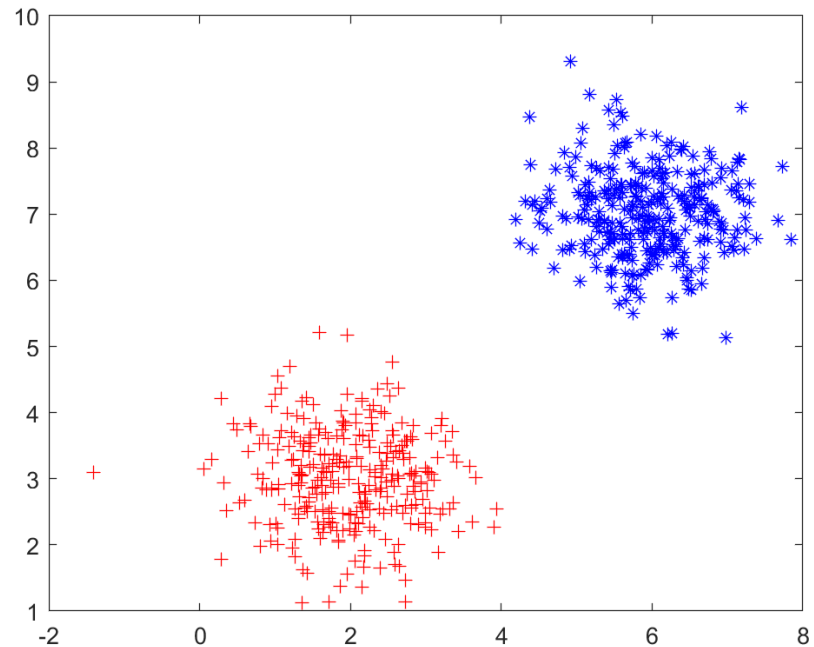
$$d = \frac{\overline{m_1} + \overline{m_2}}{2}$$

➡ 分界面方程： $(\mathbf{w}^*)^T \mathbf{x} - d = 0$



线性分类器 – Fisher线性判别

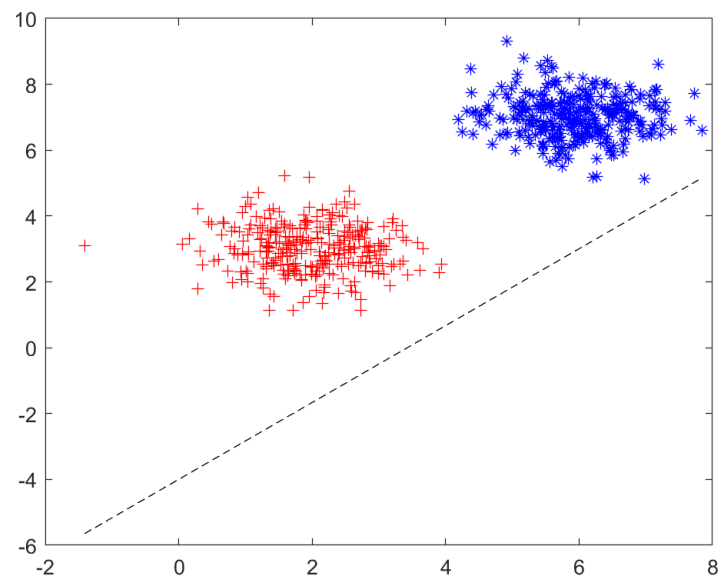
```
clc;  
close all;  
clear all;  
%% 生成数据  
randn('seed', 2020);  
mu1 = [2 3];  
sigma1 = [0.5 0;  
          0 0.5];  
data1 = mvnrnd(mu1, sigma1, 300);  
  
randn('seed', 2021);  
mu2 = [6 7];  
sigma2 = [0.5 0;  
          0 0.5];  
data2 = mvnrnd(mu2, sigma2, 300);
```





线性分类器 – Fisher线性判别

```
mu_1 = mean(data1, 1);  
mu_2 = mean(data2, 1);  
tmp = data1 - repmat(mu_1, [size(data1, 1), 1]);  
S1 = tmp' * tmp;  
tmp = data2 - repmat(mu_2, [size(data2, 1), 1]);  
S2 = tmp' * tmp;  
Sw = S1 + S2;  
% w_star = inv(Sw) * (mu_1 - mu_2)';  
w_star = Sw \ (mu_1 - mu_2)';  
%% Plot w_star  
Data = [data1; data2];  
[xmin, ymin] = min(Data, [], 1);  
[xmax, ymax] = max(Data, [], 1);  
X = xmin:0.1:xmax;  
k = w_star(2) / (w_star(1) + eps);  
plot(X, k * X - 4, 'k--');
```





线性分类器 – Fisher线性判别

%%

w_star = -w_star;

y1 = data1*w_star;

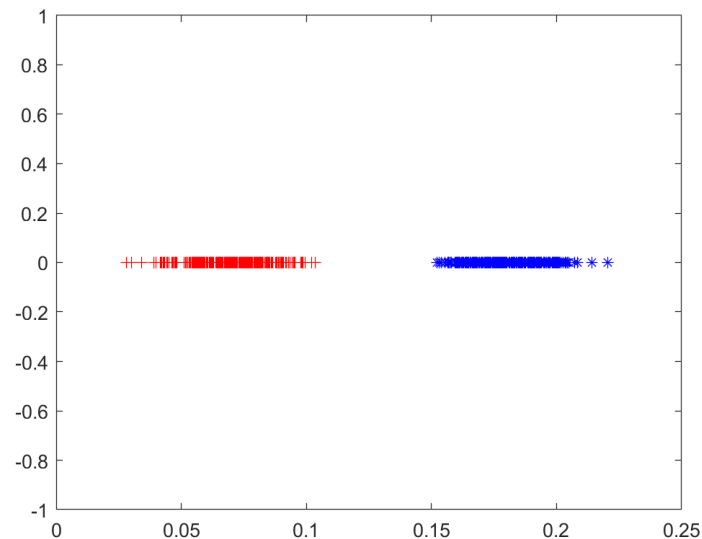
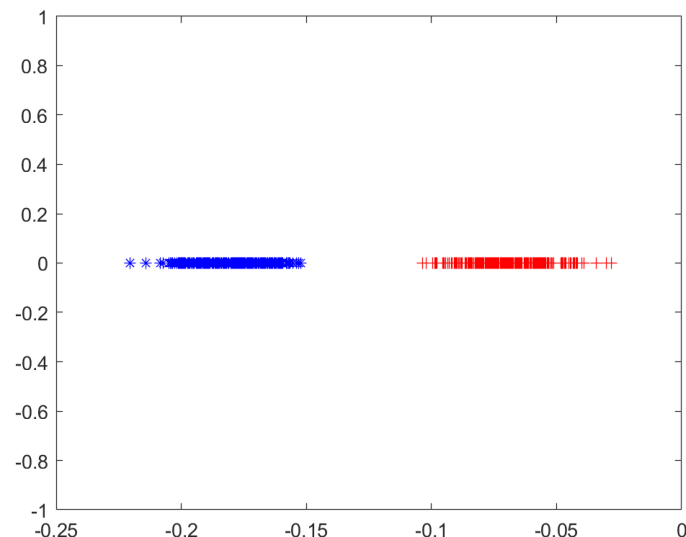
y2 = data2*w_star;

% plot((y1+4)/k, y1, 'r+');hold on;

% plot((y2+4)/k, y2, 'b*');

figure, plot(y1, zeros(length(y1)), 'r+');hold on;

plot(y2, zeros(length(y2)), 'b*');

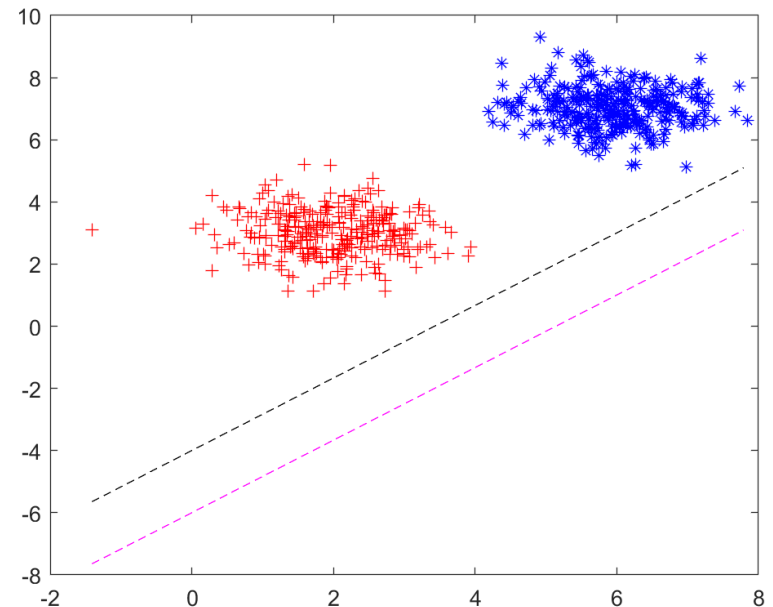




线性分类器 – Fisher线性判别

%% Sw-1Sb特征值分解

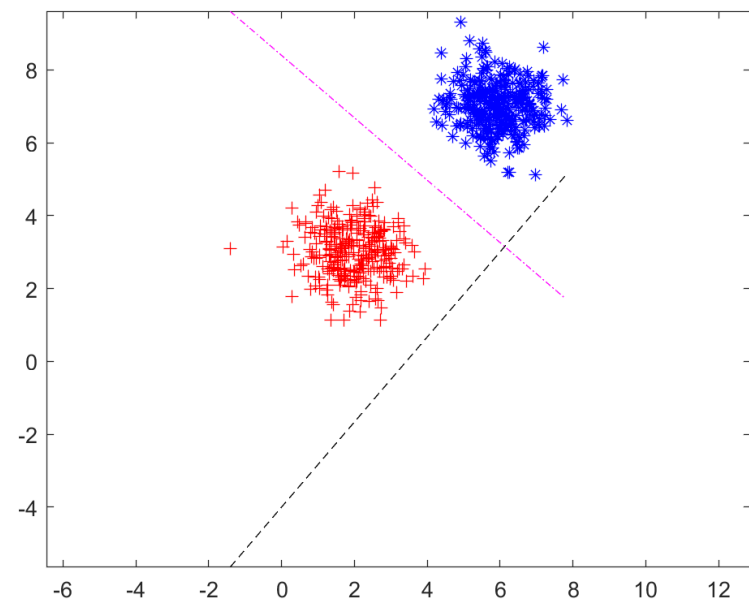
```
Sb = (mu_1-mu_2)'*(mu_1-mu_2);  
Tmp = inv(Sw)*Sb;  
[V,D] = eig(Tmp);  
D = diag(D);  
[~,ind] = max(D);  
v = V(:,ind);  
k1 = v(2)/(v(1)+eps);  
figure,plot(data1(:,1),data1(:,2),'r+');hold on;  
plot(data2(:,1),data2(:,2),'b*');hold on;  
plot(X,k*X-4,'k--');  
plot(X,k1*X-6,'m--');
```





线性分类器 – Fisher线性判别

```
%%  $G(x) = w'x + w_0$ 
mu_y1 = mean(y1);
mu_y2 = mean(y2);
d = (mu_y1 + mu_y2) / 2;
w0 = -d;
figure, plot(data1(:, 1), data1(:, 2), 'r+'); hold on;
plot(data2(:, 1), data2(:, 2), 'b*'); hold on;
plot(X, k*X-4, 'k--'); hold on;
Y = (-w_star(1)*X - w0) / (w_star(2));
plot(X, Y, 'm-');
axis equal;
```





线性分类器 – Fisher线性判别

若干素材取自网络，特此致谢！





线性分类器 – Fisher线性判别

谢谢聆听！

