

判別分析

1. 2次元の線形判別関数の傾きは最大2種類である理由がわかる
2. 線形判別関数が計算できる(2次元で3群以上分割する場合)
3. マハラビノス距離と相関係数の関係がわかる
4. マハラビノス距離から判別できる

2次元の線形判別関数の傾きは最大2種類である理由がわかる

【1】傾きの導出する2つの方法

(1) 関連記事で2次元の線形判別関数の導出を解説

まず、線形判別関数をマスターすべく、2次元の線形判別関数を導出する関連記事を確認してください。

【関連記事】線形判別関数が計算できる(2次元)

https://qcplanets.com/method/multivariate/ldf_cal2-1/

(2) 2次元の線形判別関数の傾きを求める2つの手法

関連記事にあるように2つの解法があります。

1. 相関比が最大になる条件を計算(微分、極値)
2. ラグランジュの未定乗数を使って計算(固有値)

で、解いてて気になったのは、

確かに、2つの解法とも2次方程式になるので、解は最大2種類ある。

だから、線形判別関数の傾きは最大2種類になる。

これは、納得がいくのですが、

相関比を微分すると、もともと

相関比は2次式/2次式なので、微分すると

3次式/4次式となり、

(分子)が3次式になるため解は3つになるのではないかな?

と疑問が沸きました。疑問が沸いたら、実際解いてみよう。

【2】傾きを導出する極値は最大2種類しかない

(1) 相関比関数を定義

傾きの比を $k = \frac{a_2}{a_1}$ とおくと、

(分母)、(分子)がともに2次式の相関比関数ができます。

一般化して

$$F = \frac{ba_1^2 + ca_1a_2 + da_2^2}{ea_1^2 + fa_1a_2 + ga_2^2}$$

つまり、

$$F(k) = \frac{bk^2 + ck + d}{ek^2 + fk + g}$$

として、実際に微分してみましょう。

(2) 相関比関数を微分

$$\frac{dF}{dk} = \frac{(2bk+c)(ek^2+fk+g) - (bk^2+ck+d)(2ek+f)}{(ek^2+fk+g)^2}$$

ここで、(分子)だけ取り出します。

(分子)=

$$(2bk+c)(ek^2+fk+g) - (bk^2+ck+d)(2ek+f)$$

$$= 2bek^3 + (2bf+ce)k^2 + (2bg+cf)k + cg$$

$$- (2bek^3 + (bf+2ce)k^2 + (cf+2de)k + df)$$

と3次式が出て来ますが、よく見ると

$2bek^3 - 2bek^3 = 0$ と3次項が打ち消すので、結果的に2次式になります。

よって、(分子)は

$$(分子) = (bf - ce)k^2 + 2(bg - de)k + (cg - df)$$

と2次式に収まります。

2 次式は最大 2 個の解しかできません。これで疑問が解決しました！

以上、「2 次元の線形判別関数の傾きは最大 2 種類である理由がわかる」を解説しました。

線形判別関数が計算できる(2次元で3群以上分割する場合)

【1】データを用意する

(1) 関連記事で2次元の線形判別関数の導出を解説

まず、線形判別関数をマスターすべく、2次元の線形判別関数を導出する関連記事を確認してください。

【関連記事】線形判別関数が計算できる(2次元)

https://qcplanets.com/method/multivariate/ldf_cal2-1/

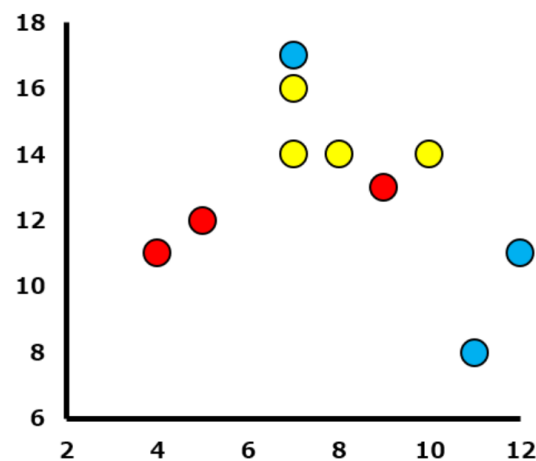
(2) データを用意する

本記事では、変数が2つ(2次元)で3群に分ける方法を解説します。

データを以下とします。

群	No	x_1	x_2	平均 $\bar{x}_1$	平均 $\bar{x}_2$
1群	1	4	11	6	12
	2	9	13		
	3	5	12		
2群	4	10	14	8	14.5
	5	7	14		
	6	8	14		
	7	7	16		
3群	8	7	17	10	12
	9	11	8		
	10	12	11		
-	-	-	全体平均	8	13

グラフで描くと下図のイメージになります。



【2】線形判別関数 $Z=0$ の条件を求める

(1) 全変動と群間変動を導出

線形判別関数 $z = a_1x_1 + a_2x_2$ と置いて、各データを下表のように表現します。

群	No	Z_i	群平均	全変動 $(Z_i - \bar{Z})^2$	群間変動 $n_k(\bar{Z}_k - \bar{Z})^2$
1群	1	$4a_1+11a_2$	$6a_1+12a_2$	$(-4a_1-2a_2)^2$	$3(-2a_1-a_2)^2$
	2	$9a_1+13a_2$		a_1^2	
	3	$5a_1+12a_2$		$(-3a_1-a_2)^2$	
2群	4	$10a_1+14a_2$	$8a_1+14.5a_2$	$(2a_1+a_2)^2$	$4(1.5a_2^2)$
	5	$7a_1+14a_2$		$(-a_1+a_2)^2$	
	6	$8a_1+14a_2$		a_2^2	
	7	$7a_1+16a_2$		$(-a_1+3a_2)^2$	
3群	8	$7a_1+17a_2$	$10a_1+12a_2$	$(-a_1+4a_2)^2$	$3(2a_1-a_2)^2$
	9	$11a_1+8a_2$		$(3a_1-5a_2)^2$	
	10	$12a_1+11a_2$		$(4a_1-2a_2)^2$	
-	平均	$8a_1+13a_2$	合計	$58a_1^2-36a_1a_2+62a_2^2$	$24a_1^2+15a_2^2$

ここで、

●全変動 $S_T=58a_1^2-36a_1a_2+62a_2^2$

●群間変動 $S_B=24a_1^2+15a_2^2$

全変動から、線形判別関数の係数 a_1, a_2 の比が計算できますが、2次元の場合は関連記事のとおり2通り解法があります。それぞれ解説します。

1. 相関比が最大になる条件を計算
2. ラグランジュの未定乗数を使って計算

(2) 相関比から傾きの比を導出

$$\text{相関比 } F = \frac{S_B}{S_T} = \frac{24a_1^2+15a_2^2}{58a_1^2-36a_1a_2+62a_2^2}$$

とおき、

$$k = \frac{a_1}{a_2} \text{ として、}$$

$$F(k) = \frac{24k^2+15}{58k^2-36k+62}$$

と変形して、この関数のグラフを描きましょう。

$F(k)$ を微分すると

$$F'(k) = \frac{-12(72k^2 - 103k - 45)}{(58k^2 - 36k + 62)^2}$$

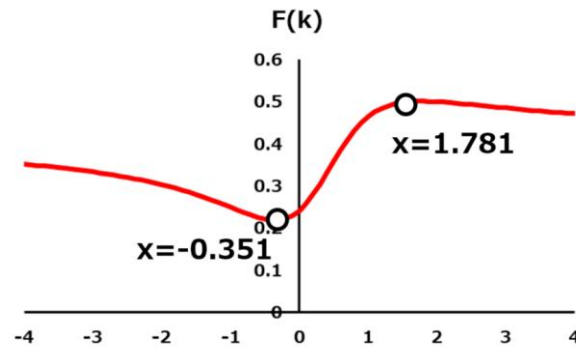
● $F'(k)=0$ は

$$72k^2 - 103k - 45 = 0 \text{より}$$

$$k = \frac{103 \pm \sqrt{23569}}{144}$$

$$= -0.351, 1.781$$

グラフを描くと下図のとおりです。



(3) ラグランジュの未定乗数法から傾きの比を導出

$58a_1^2 - 36a_1a_2 + 62a_2^2 = 1$ とした条件で、 $24a_1^2 + 15a_2^2$ の最大値を求める方法でしたね。

● $F(a_1, a_2, \lambda) = 24a_1^2 + 15a_2^2 - \lambda(58a_1^2 - 36a_1a_2 + 62a_2^2)$

として、固有方程式

● $\frac{\partial F}{\partial a_1} = 0$

● $\frac{\partial F}{\partial a_2} = 0$

を解いて、固有値解 λ 、固有ベクトルを計算します。

相関比を微分して得られた結果と一致します。(計算してみてくださいね)

【3】線形判別関数を求める

(1) 線形判別関数の傾き

結果は2つ出て来て、

$$\frac{a_1}{a_2} = -0.351, 1.781$$

ですから、

●(i) $a_1 = -0.351, a_2 = 1$

●(ii) $a_1 = 1.781, a_2 = 1$

として、線形判別関数の切片を計算しましょう。

(2) 線形判別関数の切片

ここで思うのは、

●傾きはしっかり数学するけど

●切片は帳尻合わせって感じです。

要は、

3群に分けるので、

●1つは「1群の平均」と「2群平均」の平均を通る線形判別関数

●もう1つは「2群の平均」と「3群平均」の平均を通る線形判別関数を考えます。

どこか1点を通る条件が帳尻合わせって感じなんですよ。

★平均点を求める

各群の平均はそれぞれ、

●1群： $(x, y) = (6, 12)$

●2群： $(x, y) = (8, 14.5)$

●3群： $(x, y) = (10, 12)$

なので、

●(a)「1群の平均」と「2群平均」の平均は $(x, y) = (7, 13.25)$

●(b)「2群の平均」と「3群平均」の平均は $(x, y) = (9, 13.25)$

となり、そこを通る線形判別関数を作ります。

(3)線形判別関数を求める

傾きは

●(i) $a_1 = -0.351, a_2 = 1$

●(ii) $a_1 = 1.781, a_2 = 1$

で、切片は、

●(a) $(x, y) = (7, 13.25)$

●(b) $(x, y) = (9, 13.25)$

の2×2のパターンなので、4本の線形判別関数が引けます。

結果は、下表になります。

傾きは

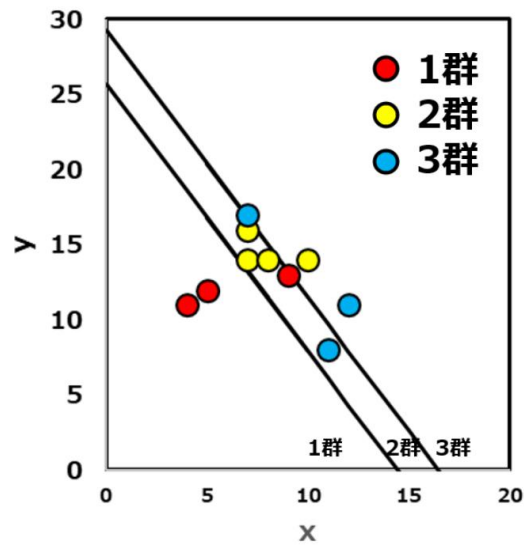
●第1判別関数 傾き1.781	●第2判別関数 傾き-0.351
$1.781x + y - 25.717 = 0$	$-0.351x + y - 10.793 = 0$
$1.781x + y - 29.279 = 0$	$-0.351x + y - 10.091 = 0$

【4】データ判別正誤率で評価する

傾きが2つあるので、それぞれの場合で正誤評価します。

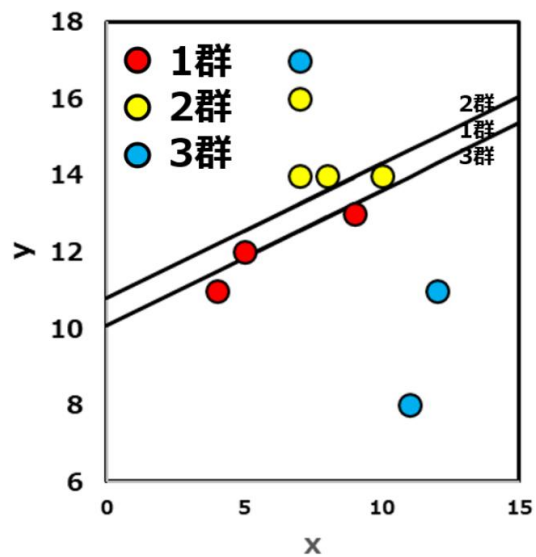
(1) 第1判別関数(傾き-0.351)

結果はグラフのとおりです。外側から1群、2群、3群とうまく区分できています。



(2) 第 2 判別関数(傾き 1.781)

結果はグラフのとおりです。外側から 2 群、1 群、3 群とうまく区分できています。



どちらも、結構うまく判別できましたね！

以上、「線形判別関数が計算できる(2次元で3群以上分割する場合)」を解説しました。

【1】マハラビノス距離を相関係数で表現する

(1) マハラノビス距離について

関連記事で復習しましょう。

【関連記事】マハラビノス距離が計算できる

https://qcplanets.com/method/multivariate/mahalanobis-distance_cal/

(2) 相関係数

相関係数は、定義通り、2変数 x_1, x_2 と平方和 S を使うと、次式で表現できます。

●相関係数 $r = \frac{S_{12}}{\sqrt{S_{11}S_{22}}}$

(3) マハラビノス距離を相関係数で表現

関連記事から、2次元のマハラノビス距離を

$$D_M^2 = (X_1, X_2) \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix}^{-1} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

と表現できますね。行列表記を実際に展開しましょう。

$$\begin{aligned} D_M^2 &= (X_1, X_2) \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix}^{-1} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= \frac{1}{S_{xx}S_{yy} - S_{xy}^2} (X_1, X_2) \begin{pmatrix} S_{22} & -S_{12} \\ -S_{12} & S_{11} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= \frac{1}{S_{11}S_{22} - S_{12}^2} (S_{22}X_1^2 - 2S_{12}X_1X_2 + S_{11}X_2^2) \\ &= (\text{式1}) \end{aligned}$$

ここで、相関係数 r を代入しましょう。

$$\begin{aligned} r &= \frac{S_{12}}{\sqrt{S_{11}S_{22}}} \text{ から} \\ r^2 &= \frac{S_{12}^2}{S_{11}S_{22}} \\ r^2 S_{11} S_{22} &= S_{12}^2 \end{aligned}$$

として、マハラノビス距離に代入します。

$$\begin{aligned} &(\text{式1}) \\ &= \frac{1}{S_{11}S_{22}(1-r^2)} X_1^2 - \frac{2r^2}{S_{12}(1-r^2)} X_1X_2 + \frac{1}{S_{11}S_{22}(1-r^2)} X_2^2 \\ &= (\text{式2}) \end{aligned}$$

$$\begin{aligned} &(\text{式2}) \\ &= \frac{X_1^2}{S_{11}(1-r^2)} - \frac{2r^2}{S_{12}(1-r^2)} X_1X_2 + \frac{X_2^2}{S_{22}(1-r^2)} \\ &= (\text{式3}) \end{aligned}$$

よって

マハラノビス距離 D_M^2 は

$$D_M^2 = \frac{X_1^2}{S_{11}(1-r^2)} - \frac{2r^2}{S_{12}(1-r^2)} X_1X_2 + \frac{X_2^2}{S_{22}(1-r^2)}$$

と表現できる

【2】 マハラビノス距離とユークリッド距離の関係

(1) ユークリッド距離

ユークリッド距離って難しい書き方ですが、単純に、
距離 $D^2 = X_1^2 + X_2^2$ です。

ただし、マハラノビス距離と比較するために、

ここでは、あえて標準化した距離をユークリッド距離として定義します。

つまり、各方向の長さの2乗を、その方向の標本分散で割ります。

$$\text{距離 } D^2 = \frac{X_1^2}{S_{11}} + \frac{X_2^2}{S_{22}}$$

(2) マハラビノス距離とユークリッド距離が等しい相関係数は0

マハラビノス距離とユークリッド距離の式を比較しましょう。

$$\bullet \text{マハラビノス距離: } D_M^2 = \frac{X_1^2}{S_{11}(1-r^2)} - \frac{2r^2}{S_{12}(1-r^2)} X_1 X_2 + \frac{X_2^2}{S_{22}(1-r^2)}$$

$$\bullet \text{ユークリッド距離: } D^2 = \frac{X_1^2}{S_{11}} + \frac{X_2^2}{S_{22}}$$

ここで、相関係数 $r=0$ をマハラビノス距離へ代入すると

$$\begin{aligned} \bullet \text{マハラビノス距離: } D_M^2 &= \frac{X_1^2}{S_{11}(1-0^2)} - \frac{2 \cdot 0^2}{S_{12}(1-0^2)} X_1 X_2 + \frac{X_2^2}{S_{22}(1-0^2)} \\ &= \frac{X_1^2}{S_{11}} + \frac{X_2^2}{S_{22}} \\ &= D^2 \end{aligned}$$

と、

相関係数 $r=0$ のときはマハラビノス距離とユークリッド距離が等しくなります。

つまり、

変数間が無相関と仮定した距離がユークリッド距離ともいえますね。

【3】 相関係数とマハラビノス距離の関係

(1) 相関係数が0の場合

先ほど解説したとおり、

相関係数 $r=0$ のときはマハラビノス距離とユークリッド距離が等しくなります。

(2) 相関係数が1に近づく場合

マハラビノス距離において、相関係数が1になると、

$$\bullet \text{マハラビノス距離: } D_M^2 = \frac{X_1^2}{S_{11}(1-r^2)} - \frac{2r^2}{S_{12}(1-r^2)} X_1 X_2 + \frac{X_2^2}{S_{22}(1-r^2)}$$

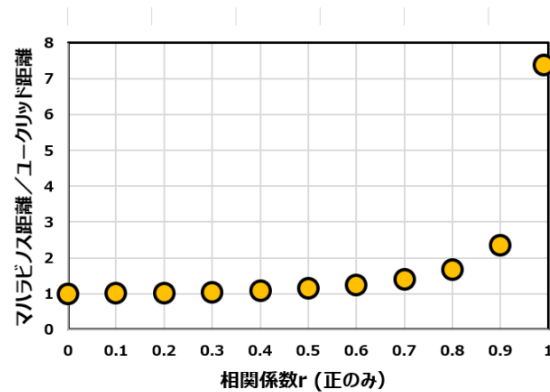
の、

分母が0に近づくので、マハラビノス距離 $\rightarrow$ 無限大に発散します。

相関係数 $r=1$ のようなほぼ直線関係のデータではマハラビノス距離は精度が悪いといえます。

(3) 相関係数とマハラビノス距離の関係

実際に、いろいろな相関係数における、マハラビノス距離とユークリッド距離の関係を調べてみると、次の結果になりました。



図における、数値はデータによって変わりますが、伝えたいことは、

相関係数 r が0に近いと、マハラビノス距離とユークリッド距離はほぼ等しいが
相関係数が増えると、マハラビノス距離がだんだん長くなり相関係数が1に近づくと発散する

面白い性質ですね。

距離といってもいろいろな距離が定義できます。
その定義した距離の特性をよく知った上で活用しましょう。

以上、「マハラビノス距離が計算できる」を解説しました。

【1】データ事例

(1) マハラノビス距離について

【関連記事】マハラビノス距離が計算できる

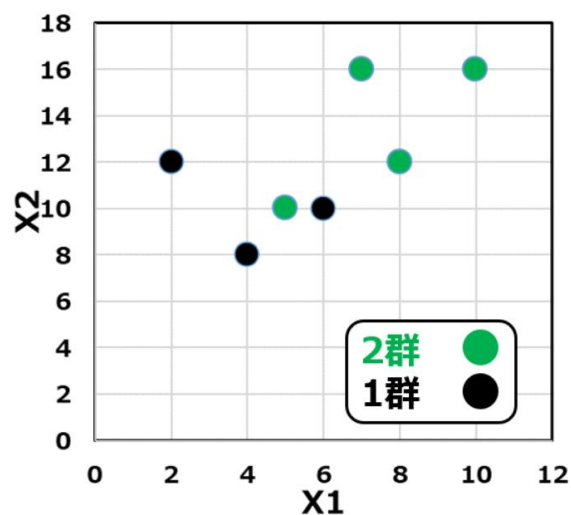
https://qcplanets.com/method/multivariate/mahalanobis-distance_cal/

(2) データ事例

下表のように、2つの群に分かれている7個のデータを用意します。

群	No	x_1	x_2	平均 $\bar{x}_1$	平均 $\bar{x}_2$
1	1	4	8	4	10
	2	6	10		
	3	2	12		
2	4	10	16	7.5	13.5
	5	5	10		
	6	8	12		
	7	7	16		
合計	42	84	全平均	6	12

グラフは下図のとおりです。



(3) 標本分散を計算

マハラビノス距離の計算に必要な標本分散を計算します。結果は下表のとおりです。

群	N0	x_1	x_2	標本平均 S_{11}	標本平均 S_{22}	標本平均 S_{12}
1	1	4	8	4	4	-2
	2	6	10			
	3	2	12			
2	4	10	16	4.33	9	4.33
	5	5	10			
	6	8	12			
	7	7	16			

【2】各群のマハラノビス距離を算出

(1) マハラノビス距離を相関係数で表現

関連記事から、2次元のマハラノビス距離を

$$D_M^2 = (X_1, X_2) \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix}^{-1} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

と表現できますね。上の式を使って、マハラノビス距離を計算します。

(2) 各群のマハラノビス距離を計算

① 1群のマハラノビス距離

$$\begin{aligned} D_M^2 &= (X_1, X_2) \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix}^{-1} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= (X_1, X_2) \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}^{-1} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= (x_1 - \bar{x}_1, x_2 - \bar{x}_2) \begin{pmatrix} 0.333 & 0.167 \\ 0.167 & 0.333 \end{pmatrix} \begin{pmatrix} x_1 - \bar{x}_1 \\ x_2 - \bar{x}_2 \end{pmatrix} \end{aligned}$$

より、結果をまとめると

$$D_M^2 = \frac{1}{3}(x_1 - 4)^2 + \frac{1}{3}(x_1 - 4)(x_2 - 10) + \frac{1}{3}(x_2 - 10)^2$$

となります。

② 2群のマハラノビス距離

$$\begin{aligned} D_M^2 &= (X_1, X_2) \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix}^{-1} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= (X_1, X_2) \begin{pmatrix} 4.33 & 4.33 \\ 4.33 & 9 \end{pmatrix}^{-1} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= (x_1 - \bar{x}_1, x_2 - \bar{x}_2) \begin{pmatrix} 0.445 & -0.214 \\ -0.214 & 0.214 \end{pmatrix} \begin{pmatrix} x_1 - \bar{x}_1 \\ x_2 - \bar{x}_2 \end{pmatrix} \end{aligned}$$

より、結果をまとめると

$$D_M^2 = 0.445(x_1 - 7.5)^2 - 0.214(x_1 - 7.5)(x_2 - 13.5) + 0.214(x_2 - 13.5)^2$$

となります。

結果をまとめると、

マハラノビス距離は

● 1群: $D_M^2 = \frac{1}{3}(x_1 - 4)^2 + \frac{1}{3}(x_1 - 4)(x_2 - 10) + \frac{1}{3}(x_2 - 10)^2$

● 2群: $D_M^2 = 0.445(x_1 - 7.5)^2 - 0.214(x_1 - 7.5)(x_2 - 13.5) + 0.214(x_2 - 13.5)^2$

各座標において、1群、2群のマハラノビス距離を計算してマハラノビス距離が短い群に属することを確認しましょう。

(3) マハラビノス距離を図示

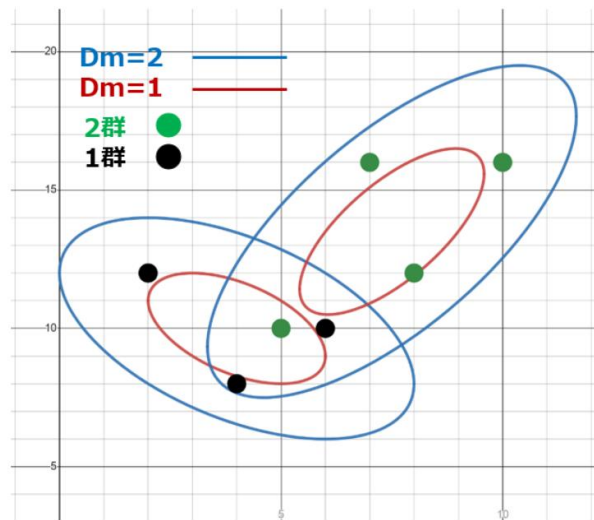
その前に、マハラノビス距離を図示しましょう。傾いた楕円形になりますね。グラフに描く式は、

図示するグラフは

● 1群: $\frac{1}{3}(x_1 - 4)^2 + \frac{1}{3}(x_1 - 4)(x_2 - 10) + \frac{1}{3}(x_2 - 10)^2 = D_M^2$

● 2群: $0.445(x_1 - 7.5)^2 - 0.214(x_1 - 7.5)(x_2 - 13.5) + 0.214(x_2 - 13.5)^2 = D_M^2$

として、 $D_M^2 = 1^2, 2^2$ について図示します。



上図の等高線を見ると、

● マハラノビス距離がどの程度あるか、

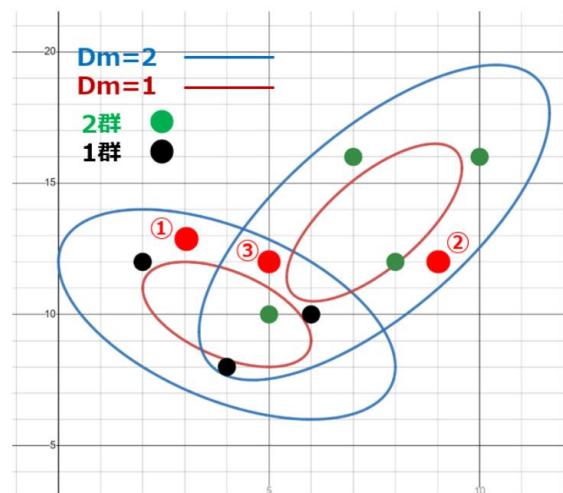
● どちらの群に属しているか、

がはっきりしますね。

【3】各点がどちらの群に属するか？

(1) 3点を例に、マハラビノス距離から判別

下図のように3つの赤い点（座標）を用意して、1群、2群どちらに属するかをみましょう。



1. $(x_1, x_2) = (3, 13)$
2. $(x_1, x_2) = (9, 12)$
3. $(x_1, x_2) = (5, 12)$

では、実際に計算すると下表のような結果になります。

データ	座標	DM1	DM2	属する群
①	(3,13)	1.53	2.85	2群
②	(9,12)	3.61	1.56	1群
③	(5,12)	1.53	1.28	2群

①②はあきらかにすぐわかりますが、

③のような、7つのデータからみると1群の3つの黒点に近い座標なので、1群に属しやすいかと思いきや、マハラノビス距離から評価すると2群に属することが分かりました。

(2) マハラノビス距離で判別するときの注意点

マハラノビス距離では、標準偏差で割ったり、相関関係を配慮した値になるので、座標から見て直観的に判別した結果と異なる場合があります。

各群のばらつき、変数（軸）どうし相関関係に注意してマハラノビス距離を計算しましょう。

その一方、線形判別関数は直線で引くので、座標からすぐ判別がつきますが、マハラノビス距離からの評価は座標をただだけでは完璧に判別できない難しさがありますね。

以上、「マハラノビス距離から判別できる」を解説しました。