



ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
AGRICULTURAL UNIVERSITY OF ATHENS

Διακριτική Ανάλυση

Κατσιλέρος Αναστάσιος

2017

ΔΙΑΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ (Discriminant Analysis)

Η ανάλυση γραμμικής διάκρισης χρησιμοποιείται ως τεχνική μείωσης των διαστάσεων για εφαρμογές ταξινόμησης ή για την καλύτερη κατανόηση των σχέσεων μπορεί να υπάρχει μεταξύ των μεταβλητών.

Δημιουργούνται διακριτικές (διαχωριστικές) συναρτήσεις :

$$Z = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_p X_p$$

και επιλέγονται αυτές που μεγιστοποιούν τις διαφορές μεταξύ των ομάδων και τις διαχώριζαν καλύτερα.

Πειραματικοί Σχεδιασμοί

```
> library(MASS)
```

```
> fit=lda(Species ~ Sepal.Length + Sepal.Width + Petal.Length + Petal.Width, data=iris)
```

```
> fit
```

```
Call:
```

```
lda(Species ~ Sepal.Length + Sepal.Width + Petal.Length + Petal.Width, data = iris)
```

```
Prior probabilities of groups:
```

```
setosa    versicolor  virginica  
0.3333333 0.3333333 0.3333333
```

```
Group means:
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
setosa	5.006	3.428	1.462	0.246
versicolor	5.936	2.770	4.260	1.326
virginica	6.588	2.974	5.552	2.026

```
Coefficients of linear discriminants:
```

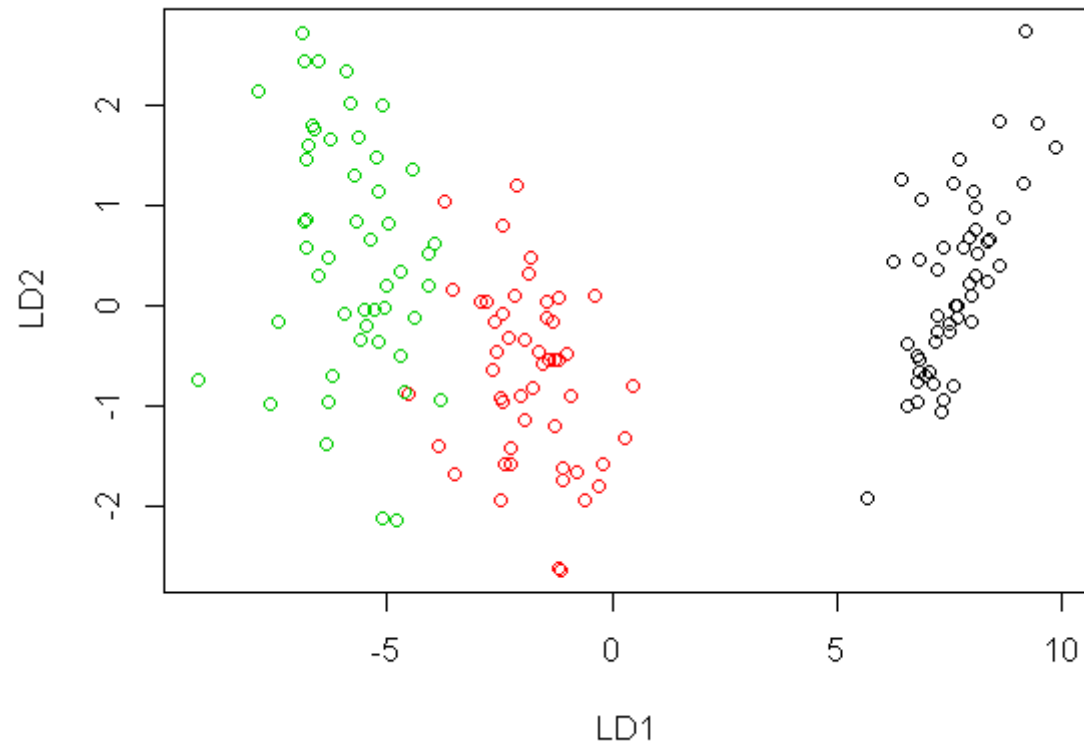
	LD1	LD2
Sepal.Length	0.8293776	0.02410215
Sepal.Width	1.5344731	2.16452123
Petal.Length	-2.2012117	-0.93192121
Petal.Width	-2.8104603	2.83918785

```
Proportion of trace:
```

LD1	LD2
0.9912	0.0088

Πειραματικοί Σχεδιασμοί

```
> plot(predict(fit)$x, col=iris$Species)
```



Πειραματικοί Σχεδιασμοί

```
> n1=length(iris[iris[,5]=="setosa",5])
> n2=length(iris[iris[,5]=="versicolor",5])
> n3=length(iris[iris[,5]=="virginica",5])
> N=n1 + n2 + n3
> Total=cov(iris[,1:4])
> w1=(n1-1)*cov(iris[iris[,5]=="setosa",1:4])
> w2=(n2-1)*cov(iris[iris[,5]=="versicolor",1:4])
> w3=(n3-1)*cov(iris[iris[,5]=="virginica",1:4])
> W=(w1+w2+w3)/(N-3)
> B= Total - W
> W # within matrix W
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length	0.26500816	0.09272109	0.16751429	0.03840136
Sepal.Width	0.09272109	0.11538776	0.05524354	0.03271020
Petal.Length	0.16751429	0.05524354	0.18518776	0.04266531
Petal.Width	0.03840136	0.03271020	0.04266531	0.04188163

```
> B # between matrix B
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length	0.4206853	-0.13515509	1.1068012	0.4778693
Sepal.Width	-0.1351551	0.07459166	-0.3848999	-0.1543496
Petal.Length	1.1068012	-0.38489991	2.9310901	1.2529441
Petal.Width	0.4778693	-0.15434958	1.2529441	0.5391246

Πειραματικοί Σχεδιασμοί

```
> eigen=eigen(solve(W) %*% B) # W-1B  
> eigen
```

```
eigen() decomposition
```

```
$values
```

```
[1] 3.219193e+01 2.853910e-01 8.118648e-15 2.136821e-15
```

```
$vectors
```

```
      [,1]      [,2]      [,3]      [,4]  
[1,] -0.2087418 -0.006531964 0.5287521 -0.8482524  
[2,] -0.3862037 -0.586610553 -0.4438915 0.1828264  
[3,] 0.5540117 0.252561540 -0.4867306 0.1409649  
[4,] 0.7073504 -0.769453092 0.5352334 0.4766248
```

Πειραματικοί Σχεδιασμοί

```
> variables=as.matrix(iris[, 1:4])  
> manova=manova(variables ~ iris$Species)
```

```
> summary(manova, test = "Hotelling-Lawley")
```

```
          Df Hotelling-Lawley approx F num Df den Df  Pr(>F)  
iris$Species 2      32.477 580.53    8 286 < 2.2e-16 ***  
Residuals 147
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> summary(manova, test = "Roy")
```

```
          Df  Roy approx F num Df den Df  Pr(>F)  
iris$Species 2 32.192 1167    4 145 < 2.2e-16 ***  
Residuals 147
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Πειραματικοί Σχεδιασμοί

```
> summary(manova, test = "Pillai")
```

```
      Df Pillai approx F num Df den Df  Pr(>F)
iris$Species  2 1.1919  53.466   8  290 < 2.2e-16 ***
Residuals    147
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> summary(manova, test = "Wilks")
```

```
      Df  Wilks approx F num Df den Df  Pr(>F)
iris$Species  2 0.023439 199.15   8  288 < 2.2e-16 ***
Residuals    147
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```