Discriminant analysis

- ANOVA and MANOVA: predict a (counted/measured) response from group membership.
- Discriminant analysis: predict group membership based on counted/measured variables.
- Covers same ground as logistic regression (and its variations), but emphasis on classifying observed data into correct groups.
- Does so by searching for linear combination of original variables that best separates data into groups (canonical variables).
- Assumption here that groups are known (for data we have). If trying to “best separate” data into unknown groups, see *cluster analysis*.
- Examples: revisit seed yield and weight data, professions/activities data; remote-sensing data.

Example 1: seed yields and weights

Recall data from MANOVA: needed a multivariate analysis to find difference in seed yield and weight based on whether they were high or low fertilizer.

```
   Plot of yield*weight$fertilizer.  Symbol points to label.
yield
   38 +                            > high
      |
   37 +
      |
   36 +
      |
   35 +               > low                             high <
      |
   34 +  > low                                  > high
      |
   33 +                                                 high <
      |
   32 +                                         > low
      |
   31 +
      |
   30 +
      |
   29 +                                                  low <
      ---+------------+------------+------------+------------+--
        10           11           12           13           14
                              weight
```

Plot variations

Above plot produced with
`plot yield * weight $ fertilizer`. Compare
`plot yield * weight = fertilizer`:

```
          Plot of yield*weight.  Symbol is value of fertilizer.
     yield
        38 +                             h
           |
        37 +
           |
        36 +
           |
        35 +              l                          h
           |
        34 +  l                                h
           |
        33 +                                         h
           |
        32 +                                   l
           |
        31 +
           |
        30 +
           |
        29 +                                         l
           ---+------------+------------+------------+------------+--
             10           11           12           13           14

                                 weight
```

Basic PROC DISCRIM

We found it was a *combination* of weight and yield that distinguished high from low fertilizer.

```
data manova1;
  infile "manova1.dat";
  input fertilizer $ yield weight;
proc discrim can list out=x;
  class fertilizer;
  var yield weight;
```

In PROC DISCRIM:

- `can` gets “canonical variables analysis”
- `list` lists observations and summarizes classification
- output data set gives “canonical variable scores” for each observation

Don’t need both `list` and output data set; choose according to needs.

Output

The DISCRIM Procedure

Observations	8	DF Total	7
Variables	2	DF Within Classes	6
Classes	2	DF Between Classes	1

Class Level Information

fertilizer	Variable Name	Frequency	Weight	Proportion	Prior Probability
high	high	4	4.0000	0.500000	0.500000
low	low	4	4.0000	0.500000	0.500000

Summarizes input: 8 observations, 2 classes (high and low), 4 observations in each class.

More output

```
Test of H0: The canonical correlations in the
   current row and all that follow are zero

     Likelihood    Approximate
          Ratio        F Value    Num DF    Den DF    Pr > F

1    0.19845779          10.10         2         5    0.0175

           NOTE: The F statistic is exact.
```

That is, we really do have $1 + 1 = 2$ groups (the “highs” and “lows” are not all mixed up).

Canonical coefficients

```
Raw Canonical Coefficients

Variable                  Can1

yield              0.766676064
weight             1.251356335

Class Means on Canonical Variables

fertilizer                 Can1

high                1.740442790
low                -1.740442790
```

The combination $0.77 yield + 1.25 weight$ best separates the highs from the lows. When you do this (and standardize the results: see below) a positive value of Can1 goes with “high” and a negative goes with “low”.

Output from “list”

Posterior Probability of Membership in fertilizer

Obs	From fertilizer	Classified into fertilizer	high	low
1	low	low	0.0000	1.0000
2	low	low	0.0012	0.9988
3	low	low	0.0232	0.9768
4	low	low	0.0458	0.9542
5	high	high	0.9818	0.0182
6	high	high	0.9998	0.0002
7	high	high	0.9089	0.0911
8	high	high	0.9999	0.0001

Summary of estimated probabilities that observation with those values of seed yield and seed weight would be classified into each fertilizer category. See that each classification was correct, emphasized below:

Classification summary

Number of Observations and Percent Classified into fertilizer

From fertilizer	high	low	Total
high	4	0	4
	100.00	0.00	100.00
low	0	4	4
	0.00	100.00	100.00
Total	4	4	8
	50.00	50.00	100.00
Priors	0.5	0.5	

Error Count Estimates for fertilizer

	high	low	Total
Rate	0.0000	0.0000	0.0000
Priors	0.5000	0.5000	

Output data set

Finally, the output data set, like the output from `list`, but with more detail:

Obs	fertilizer	yield	weight	Can1	Can2	high	low	_INTO_
1	low	34	10	-3.09314	.	0.00002	0.99998	low
2	low	29	14	-1.92110	.	0.00125	0.99875	low
3	low	35	11	-1.07511	.	0.02315	0.97685	low
4	low	32	13	-0.87242	.	0.04579	0.95421	low
5	high	33	14	1.14561	.	0.98180	0.01820	high
6	high	38	12	2.47628	.	0.99982	0.00018	high
7	high	34	13	0.66093	.	0.90893	0.09107	high
8	high	35	14	2.67896	.	0.99991	0.00009	high

Shows original variable values plus scores on first canonical variable (the one that best separates observations into correct categories). Here `Can1` scaled to have mean 0 (overall) and SD 1 for each group.

Example 2: professions and leisure activities

- Same data we used for profile analysis (some):

```
bellydancer 7 10 6 5
bellydancer 8 9 5 7
bellydancer 5 10 5 8
politician 5 5 5 6
politician 4 5 6 5
admin 4 2 2 5
admin 7 1 2 4
admin 6 3 3 3
```

- How can we best use the scores on the activities to predict a person’s profession?
- Or, what combination(s) of scores best separate data into profession groups?

Some SAS code

```
data profile;
  infile "profile.dat";
  input group $ read dance tv ski;

proc discrim can list out=fred;
  class group;

proc print data=fred;

proc plot data=fred;
  plot Can1 * Can2 = group;
```

Can also specify `read`, `dance`, `tv` and `ski` on a `var` line in PROC DISCRIM; by default all other variables used. (Same idea as PROC MEANS.)

Obtain output data set and plot 1st 2 canonical variables.

Some output

The DISCRIM Procedure

Total Sample Size	15	DF Total	14
Variables	4	DF Within Classes	12
Classes	3	DF Between Classes	2

Number of Observations Read	15
Number of Observations Used	15

Class Level Information

group	Variable Name	Frequency	Weight	Proportion	Prior Probability
admin	admin	5	5.0000	0.333333	0.333333
bellydan	bellydan	5	5.0000	0.333333	0.333333
politici	politici	5	5.0000	0.333333	0.333333

Distances between groups

Generalized Squared Distance to group

From group	admin	bellydan	politici
admin	0	77.68532	25.14460
bellydan	77.68532	0	27.90946
politici	25.14460	27.90946	0

Bellydancers are very different overall from administrators.

Eigenvalues of Inv(E)*H
= CanRsq/(1-CanRsq)

	Eigenvalue	Difference	Proportion	Cumulative
1	16.1922	14.2262	0.8917	0.8917
2	1.9660		0.1083	1.0000

2 eigenvalues (it takes 2 lines to divide data into 3 groups), but 1st much bigger than 2nd, so data close to 1-dimensional (see on graph later).

How many canonical variables do I need?

Next table shows this:

Test of H0: The canonical correlations in the current row and all that follow are zero

	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.01961069	13.82	8	18	<.0001
2	0.33715124	6.55	3	10	0.0100

- 1st row says “need at least 1”; 2nd row says “need at least 2”.
- Max number of canonical variables is smaller of:
  - number of variables used to assess grouping (4 here)
  - number of groups minus 1 ($3 - 1 = 2$).
- Why: with g groups, $g - 1$ variables separate into that many groups.

What separates the groups

Look at “raw canonical coefficients”:

Raw Canonical Coefficients

Variable	Can1	Can2
read	0.012974652	-0.474808056
dance	0.952123961	-0.461497594
tv	0.474172636	1.244632708
ski	-0.041536839	-0.203312237

- 1st canonical variable is mostly attitudes towards dance, with a small amount of attitudes towards TV.
- 2nd is attitudes towards TV-watching contrasted with everything else.
- Bellydancers loved dancing, so Can1 distinguishes them.
- Administrators and bellydancers both hated TV compared to everything else, while politicians indifferent. (Can2 distinguishes politicians.)

Output from “list”

. . . shows that groups are pretty separate:

Posterior Probability of Membership in group

Obs	From group	Classified into group	admin	bellydan	politici
1	bellydan	bellydan	0.0000	1.0000	0.0000
2	bellydan	bellydan	0.0000	1.0000	0.0000
3	bellydan	bellydan	0.0000	1.0000	0.0000
4	bellydan	bellydan	0.0000	1.0000	0.0000
5	bellydan	bellydan	0.0000	0.9973	0.0027
6	politici	politici	0.0028	0.0000	0.9972
7	politici	politici	0.0001	0.0000	0.9999
8	politici	politici	0.0000	0.0000	1.0000
9	politici	politici	0.0000	0.0021	0.9979
10	politici	politici	0.0000	0.0000	1.0000
11	admin	admin	1.0000	0.0000	0.0000
12	admin	admin	1.0000	0.0000	0.0000
13	admin	admin	1.0000	0.0000	0.0000
14	admin	admin	1.0000	0.0000	0.0000
15	admin	admin	0.9821	0.0000	0.0179

Classification summary

shows that everyone got classified into the right job:

Number of Observations and Percent Classified into group

From group	admin	bellydan	politici	Total
admin	5	0	0	5
	100.00	0.00	0.00	100.00
bellydan	0	5	0	5
	0.00	100.00	0.00	100.00
politici	0	0	5	5
	0.00	0.00	100.00	100.00
Total	5	5	5	15
	33.33	33.33	33.33	100.00
Priors	0.33333	0.33333	0.33333	

Output data set

contains a bit more detail (note column names *vertical*):

```
                                                      b       p
                                                      e       o
                                                      l       l       _
      g        d                              a       l       i       I
      r     r  a         C        C    C C    d       y       t       N
 O    o     e  n   s     a        a    a a    m       d       i       T
 b    u     a  c t k     n        n    n n    i       a       c       O
 s    p     d  e v i     1        2    3 4    n       n       i       _

 1 bellydan 7 10 6 5  5.23731 -0.58059 . . 0.00000 1.00000 0.00000 bellydan
 2 bellydan 8  9 5 7  3.74092 -2.24515 . . 0.00000 1.00000 0.00000 bellydan
 3 bellydan 5 10 5 8  4.61258 -1.48554 . . 0.00000 1.00000 0.00000 bellydan
 4 bellydan 6 10 6 8  5.09973 -0.71571 . . 0.00000 1.00000 0.00000 bellydan
 5 bellydan 7  8 7 9  3.64109  0.77379 . . 0.00000 0.99729 0.00271 bellydan
 6 politici 4  4 4 4 -1.42116  1.32687 . . 0.00283 0.00000 0.99717 politici
 7 politici 6  4 5 3 -0.87950  1.82520 . . 0.00008 0.00000 0.99992 politici
 8 politici 5  5 5 6 -0.06496  1.22857 . . 0.00001 0.00000 0.99998 politici
 9 politici 6  6 6 7  1.33277  1.33359 . . 0.00000 0.00214 0.99786 politici
10 politici 4  5 6 5  0.43777  3.15133 . . 0.00000 0.00000 1.00000 politici
11 admin    3  1 1 2 -5.62995 -0.14110 . . 1.00000 0.00000 0.00000 admin
12 admin    5  3 1 5 -3.82437 -2.62365 . . 1.00000 0.00000 0.00000 admin
13 admin    4  2 2 5 -4.31529 -0.44271 . . 0.99999 0.00000 0.00001 admin
14 admin    7  1 2 4 -5.18696 -1.20233 . . 1.00000 0.00000 0.00000 admin
15 admin    6  3 3 3 -2.77997 -0.20257 . . 0.98209 0.00000 0.01791 admin
```

Plotting 1st 2 canonical variables

```
               Plot of Can1*Can2.  Symbol is value of group.

   5 +                b      bb
     |
     |         b                          b
     |
     |                                         p
     |                                                         p
   0 +                                        p
     |                                             p
     |                                         p
Can1 |                           a
     |
     |     a                   a
  -5 +                  a
     |                            a
     |
     |
     |
     |
 -10 +
     ---+--------+--------+--------+--------+--------+--------+--------+--
       -3       -2       -1        0        1        2        3        4

                                      Can2
```

Comments

- Even though had 4 variables, can plot 1st 2 canonical variables to “see” data. True regardless of number of original variables (though won’t see everything if more canonical variables useful).
- See that Can1 separates bellydancers (b) from administrators (a); Can2 separates politicians (p) from rest, and clarifies the position of politicians relative to others.

What if groups aren’t all distinct?

```
                  Plot of x*y.  Symbol is value of group.
x
7 +                            c             b
  |
  |
6 +                                          b             c
  |
  |
5 +                                                                      b
  |
  |
4 +                            b             c                           c
  |
  |
3 +
  |
  |
2 +              a
  |
  |
1 +a                           a             a
  -+-------------+-------------+-------------+-------------+-------------+-
   2             3             4             5             6             7

                                      y
```

SAS code

```
data mix;
  infile "mixup.dat";
  input group $ x y;
proc discrim can list out=xx;
  class group;
  var x y;
proc print;
proc plot;
  plot Can1 * Can2 = group;
```

Original data has 2 variables (x and y), so can be plotted.
Perform discriminant analysis with output data set, and plot 1st 2 canonical variables.

Distances

Generalized Squared Distance to group

From group	a	b	c
a	0	18.65441	17.88235
b	18.65441	0	0.06618
c	17.88235	0.06618	0

Groups b and c could be hard to tell apart.

Just one useful canonical variable

Eigenvalues of Inv(E)*H
= CanRsq/(1-CanRsq)

	Eigenvalue	Difference	Proportion	Cumulative
1	5.4098	5.3969	0.9976	0.9976
2	0.0129		0.0024	1.0000

Test of H0: The canonical correlations in the
current row and all that follow are zero

	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.15402685	6.19	4	16	0.0033
2	0.98727677	0.12	1	9	0.7412

With 2 variables, can only be max 2, but smallness of eigenvalue and non-significance of test tell us 2nd is not useful.

One variable *might* separate all 3 groups, however.

Canonical variables

Raw Canonical Coefficients

Variable	Can1	Can2
x	0.8252532609	-.3003312927
y	0.4629576531	0.6627706863

1st one is combination of x and y, x weighted more heavily.

Class Means on Canonical Variables

group	Can1	Can2
a	-2.848143534	-0.002552303
b	1.469358718	-0.119111596
c	1.378784816	0.121663899

Can1 separates group a from rest, Can2 doesn't do much of anything. Neither distinguishes groups b and c.

Classification

Posterior Probability of Membership in group

Obs	From group	Classified into group		a	b	c
1	a	a		1.0000	0.0000	0.0000
2	a	a		0.9982	0.0006	0.0012
3	a	a		0.9989	0.0005	0.0006
4	a	a		0.9998	0.0001	0.0002
5	b	c	*	0.0000	0.4387	0.5613
6	b	c	*	0.0961	0.4428	0.4611
7	b	b		0.0000	0.5703	0.4297
8	b	b		0.0000	0.5339	0.4660
9	c	b	*	0.0000	0.5046	0.4954
10	c	b	*	0.0000	0.5989	0.4011
11	c	c		0.0003	0.4028	0.5969
12	c	c		0.0144	0.4539	0.5317

* Misclassified observation

The a’s are very clear, but even when b’s and c’s are correctly classified, it’s a very close call.

Classification summary

Doesn't look *so* bad, but overall a third of the 12 observations wrongly classified (and doesn't show how close a call it was).

Number of Observations and Percent Classified into group

From group	a	b	c	Total
a	4	0	0	4
	100.00	0.00	0.00	100.00
b	0	2	2	4
	0.00	50.00	50.00	100.00
c	0	2	2	4
	0.00	50.00	50.00	100.00
Total	4	4	4	12
	33.33	33.33	33.33	100.00

Error Count Estimates for group

	a	b	c	Total
Rate	0.0000	0.5000	0.5000	0.3333

Canonical variable plot

```
               Plot of Can1*Can2.  Symbol is value of group.
Can1 |
   4 +
     |
     |
     |                  b
   2 +   c                    b              c                    b
     |
     |                                                                   c
     |
   0 +                                      c
     |                       b
     |
     |
  -2 +
     |                      a                                  a
     |                                           a
     |              a
  -4 +
     |
     --+----------+----------+----------+----------+----------+----------+-
     -1.5       -1.0       -0.5        0.0        0.5        1.0        1.5

                                     Can2
```

Example 3: remote-sensing data

- View 38 crops from air, measure 4 variables `x1-x4`.
- Go back and record what each crop was.
- Can we use the 4 variables to distinguish crops?
- Two new things:
  - (Linear) discriminant analysis assumes "equal covariance matrices", loosely each group has same spread and correlations between all variables. Assumed so far. Can be tested, and if fails, can do *quadratic discriminant analysis*.
  - Using same data to develop discrimination *and* to test performance is optimistic; may not generalize to other data. *Cross-validation* more honest: sees how each observation's group predicted from discriminant analysis based on *rest* of data.
  - SAS can do these. "pooled=yes" means "do linear", "pooled=no" means "do quadratic", "pooled=test" means "do test and do appropriate one". "Crosslist" option means produce classification by cross-validation.

The resulting SAS code

```
options linesize=75;

   data crops;
      infile "remote-sensing.dat";
      input Crop $ x1-x4 label $;

   proc discrim can list pool=test out=zz crosslist;
      class Crop;
      var x1-x4;

   proc plot vpercent=50;
      plot Can1 * Can2 = label;
```

Some crop names begin with same letter, so include distinct labels in data file for plotting of canonical variables.

Summary of data

The DISCRIM Procedure

Observations	36	DF Total	35
Variables	4	DF Within Classes	31
Classes	5	DF Between Classes	4

Class Level Information

Crop	Variable Name	Frequency	Weight	Proportion	Prior Probability
Clover	Clover	11	11.0000	0.305556	0.200000
Corn	Corn	7	7.0000	0.194444	0.200000
Cotton	Cotton	6	6.0000	0.166667	0.200000
Soybeans	Soybeans	6	6.0000	0.166667	0.200000
Sugarbee	Sugarbee	6	6.0000	0.166667	0.200000

36 crops, of which 11 (31%) are clover.

Assessing equality of covariance matrices

Within Covariance Matrix Information

Crop	Covariance Matrix Rank	Natural Log of the Determinant of the Covariance Matrix
Clover	4	23.64618
Corn	4	11.13472
Cotton	4	13.23569
Soybeans	4	12.45263
Sugarbee	4	17.76293
Pooled	4	21.30189

If (population) covariance matrices equal, last column should be roughly constant: not plausible here. Formal test:

Chi-Square	DF	Pr > ChiSq
98.022966	40	<.0001

Covariance matrices not equal. So use separate covariance matrices for each crop. (SAS decides with $\alpha = 0.10$).

How distinct are the groups?

Generalized Squared Distance to Crop

From Crop	Clover	Corn	Cotton	Soybeans	Sugarbee
Clover	23.64618	1317.00000	100.59945	190.52195	27.82464
Corn	25.36684	11.13472	146.92411	34.77900	21.97069
Cotton	24.01420	585.58710	13.23569	48.44914	33.57208
Soybeans	24.70009	43.14609	37.43279	12.45263	19.57568
Sugarbee	24.43063	328.84042	40.39929	104.37324	17.76293

Only some pairs of groups look at all easy to distinguish.

How many canonical variables?

	Eigenvalue	Difference	Proportion	Cumulative
1	0.6742	0.4925	0.7364	0.7364
2	0.1817	0.1289	0.1985	0.9349
3	0.0528	0.0459	0.0576	0.9925
4	0.0068		0.0075	1.0000

Test of H0: The canonical correlations in the current row and all that follow are zero

	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.47687044	1.48	16	86.179	0.1271
2	0.79837318	0.76	9	70.729	0.6515
3	0.94343017	0.44	4	60	0.7769
4	0.99319917	0.21	1	31	0.6482

4th one has very small eigenvalue: contributes nothing. Indeed, not even first significant. (Look nonetheless at plot of first two.)

Crop means on canonical variables

Class Means on Canonical Variables

Crop	Can1	Can2	Can3	Can4
Clover	0.897881914	0.171142956	-0.159468473	-0.028427125
Corn	-1.154423506	0.297279119	-0.011822020	-0.086854272
Cotton	0.155788168	0.379410840	0.348614473	0.089639679
Soybeans	-0.629213609	-0.299565534	-0.248541709	0.118577501
Sugarbee	0.174136022	-0.740433032	0.206078461	-0.054770800

Can1 distinguishes clover from corn and maybe soybeans.
Can2, if anything, picks out sugarbeet.

Classification

Posterior Probability of Membership in Crop

Obs	From Crop	Classified into Crop	Clover	Corn	Cotton	Soybeans	Sugarbee
1	Corn	Corn	0.0097	0.9810	0.0000	0.0000	0.0093
2	Corn	Corn	0.0010	0.9946	0.0000	0.0000	0.0045
3	Corn	Corn	0.0015	0.9809	0.0000	0.0000	0.0177
4	Corn	Corn	0.0068	0.9815	0.0000	0.0024	0.0093
5	Corn	Corn	0.0039	0.9835	0.0000	0.0000	0.0126
6	Corn	Corn	0.0044	0.9424	0.0000	0.0000	0.0532
7	Corn	Corn	0.0008	0.9992	0.0000	0.0000	0.0000
8	Soybeans	Soybeans	0.0053	0.0033	0.0000	0.9821	0.0092
9	Soybeans	Soybeans	0.0143	0.0000	0.0014	0.7647	0.2196
10	Soybeans	Soybeans	0.0034	0.0000	0.0002	0.9896	0.0068
11	Soybeans	Soybeans	0.0058	0.0000	0.0000	0.9854	0.0088
12	Soybeans	Soybeans	0.0072	0.0000	0.0000	0.9921	0.0007
13	Soybeans	Soybeans	0.0149	0.0000	0.0000	0.9850	0.0001
14	Cotton	Cotton	0.0157	0.0000	0.9718	0.0032	0.0093
15	Cotton	Cotton	0.0198	0.0000	0.7925	0.0004	0.1873
16	Cotton	Cotton	0.0290	0.0000	0.9590	0.0000	0.0120
17	Cotton	Cotton	0.0067	0.0000	0.9407	0.0446	0.0080
18	Cotton	Cotton	0.0051	0.0000	0.9949	0.0000	0.0000
19	Cotton	Cotton	0.0024	0.0000	0.9976	0.0000	0.0000

The rest

Obs	From Crop	Classified into Crop		Clover	Corn	Cotton	Soybeans	Sugarbee
20	Sugarbee	Soybeans	*	0.0255	0.0000	0.0000	0.8227	0.1518
21	Sugarbee	Cotton	*	0.0112	0.0000	0.5014	0.4366	0.0507
22	Sugarbee	Sugarbee		0.0422	0.0000	0.0000	0.0000	0.9578
23	Sugarbee	Sugarbee		0.1705	0.0000	0.0000	0.0000	0.8295
24	Sugarbee	Sugarbee		0.1207	0.0000	0.0000	0.0131	0.8663
25	Sugarbee	Sugarbee		0.0052	0.0000	0.0000	0.0000	0.9948
26	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
27	Clover	Clover		0.9470	0.0000	0.0000	0.0001	0.0529
28	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
29	Clover	Clover		0.9790	0.0000	0.0000	0.0000	0.0210
30	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
31	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
32	Clover	Sugarbee	*	0.1612	0.0000	0.0000	0.0000	0.8388
33	Clover	Sugarbee	*	0.1885	0.0000	0.0000	0.0000	0.8115
34	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
35	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
36	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000

Only 4 crops misclassified.

Misclassification summary

Number of Observations and Percent Classified into Crop

From Crop	Clover	Corn	Cotton	Soybeans	Sugarbee	Total
Clover	9	0	0	0	2	11
	81.82	0.00	0.00	0.00	18.18	100.00
Corn	0	7	0	0	0	7
	0.00	100.00	0.00	0.00	0.00	100.00
Cotton	0	0	6	0	0	6
	0.00	0.00	100.00	0.00	0.00	100.00
Soybeans	0	0	0	6	0	6
	0.00	0.00	0.00	100.00	0.00	100.00
Sugarbee	0	0	1	1	4	6
	0.00	0.00	16.67	16.67	66.67	100.00
Total	9	7	7	7	6	36
	25.00	19.44	19.44	19.44	16.67	100.00

Error Count Estimates for Crop

	Clover	Corn	Cotton	Soybeans	Sugarbee	Total
Rate	0.1818	0.0000	0.0000	0.0000	0.3333	0.1030

2 clover were classified as sugarbeet; 2 sugarbeet were classified as something else.

Cross-validation results are quite different

Posterior Probability of Membership in Crop

Obs	From Crop	Classified into Crop		Clover	Corn	Cotton	Soybeans	Sugarbee
1	Corn	Clover	*	0.5114	0.0000	0.0000	0.0000	0.4886
2	Corn	Corn		0.0014	0.9921	0.0000	0.0000	0.0065
3	Corn	Corn		0.0023	0.9699	0.0000	0.0000	0.0277
4	Corn	Sugarbee	*	0.3692	0.0000	0.0000	0.1291	0.5017
5	Corn	Sugarbee	*	0.2362	0.0004	0.0000	0.0000	0.7634
6	Corn	Sugarbee	*	0.0753	0.0190	0.0000	0.0000	0.9057
7	Corn	Clover	*	0.9998	0.0000	0.0000	0.0000	0.0002
8	Soybeans	Soybeans		0.0257	0.0161	0.0000	0.9136	0.0446
9	Soybeans	Sugarbee	*	0.0606	0.0000	0.0059	0.0000	0.9334
10	Soybeans	Soybeans		0.0065	0.0000	0.0003	0.9803	0.0129
11	Soybeans	Sugarbee	*	0.3965	0.0000	0.0000	0.0000	0.6035
12	Soybeans	Clover	*	0.9171	0.0000	0.0000	0.0000	0.0829
13	Soybeans	Clover	*	0.9944	0.0000	0.0000	0.0000	0.0056
14	Cotton	Cotton		0.1428	0.0000	0.7439	0.0291	0.0842
15	Cotton	Sugarbee	*	0.0954	0.0000	0.0000	0.0021	0.9025
16	Cotton	Clover	*	0.7066	0.0000	0.0000	0.0000	0.2934
17	Cotton	Cotton		0.0159	0.0000	0.8595	0.1056	0.0190
18	Cotton	Clover	*	1.0000	0.0000	0.0000	0.0000	0.0000
19	Cotton	Clover	*	1.0000	0.0000	0.0000	0.0000	0.0000

The rest

Obs	From Crop	Classified into Crop		Clover	Corn	Cotton	Soybeans	Sugarbee
20	Sugarbee	Soybeans	*	0.0300	0.0000	0.0000	0.9700	0.0000
21	Sugarbee	Cotton	*	0.0118	0.0000	0.5282	0.4599	0.0000
22	Sugarbee	Sugarbee		0.0694	0.0000	0.0000	0.0000	0.9306
23	Sugarbee	Clover	*	1.0000	0.0000	0.0000	0.0000	0.0000
24	Sugarbee	Clover	*	0.9023	0.0000	0.0000	0.0977	0.0000
25	Sugarbee	Clover	*	1.0000	0.0000	0.0000	0.0000	0.0000
26	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
27	Clover	Clover		0.5477	0.0000	0.0000	0.0008	0.4514
28	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
29	Clover	Clover		0.9694	0.0000	0.0000	0.0000	0.0306
30	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
31	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
32	Clover	Sugarbee	*	0.0441	0.0000	0.0000	0.0000	0.9559
33	Clover	Sugarbee	*	0.1352	0.0000	0.0000	0.0000	0.8648
34	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
35	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000
36	Clover	Clover		1.0000	0.0000	0.0000	0.0000	0.0000

A lot of misclassifications, and in some cases the estimated probabilities are quite low.

Cross-validation misclassification error summary

Number of Observations and Percent Classified into Crop

From Crop	Clover	Corn	Cotton	Soybeans	Sugarbee	Total
Clover	9	0	0	0	2	11
	81.82	0.00	0.00	0.00	18.18	100.00
Corn	2	2	0	0	3	7
	28.57	28.57	0.00	0.00	42.86	100.00
Cotton	3	0	2	0	1	6
	50.00	0.00	33.33	0.00	16.67	100.00
Soybeans	2	0	0	2	2	6
	33.33	0.00	0.00	33.33	33.33	100.00
Sugarbee	3	0	1	1	1	6
	50.00	0.00	16.67	16.67	16.67	100.00
Total	19	2	3	3	9	36
	52.78	5.56	8.33	8.33	25.00	100.00

Error Count Estimates for Crop

	Clover	Corn	Cotton	Soybeans	Sugarbee	Total
Rate	0.1818	0.7143	0.6667	0.6667	0.8333	0.6126

A whopping 61% of the crops are misclassified this more honest way. Sugarbeet was especially hard to get right.

Plot of 1st 2 canonical variables

Perhaps surprising that *any* method got much right!

```
               Plot of Can1*Can2.  Symbol is value of label.

   4 +                                             l
     |                                           l
     |
     |
     |
     |
   2 +
     |
     |                              lg
Can1 |                         l                                    t
     |  g
     |              l   ly       g     tt
   0 +                            g       t
     |                              tg t      l                l l
     |                              y yy  gy
     |                                  rr  r
     |                                 y   r r    r
     |
  -2 +
     --+----------+----------+----------+----------+----------+----------+-
      -3         -2         -1          0          1          2          3

                                      Can2
```

Can1 distinguishes Corn (r) and sometimes Clover (l).