The remote-sensing data

March 8, 2011

The data:

```
Corn       16 27 31 33 r
Corn       15 23 30 30 r
Corn       16 27 27 26 r
Corn       18 20 25 23 r
Corn       15 15 31 32 r
Corn       15 32 32 15 r
Corn       12 15 16 73 r
Soybeans   20 23 23 25 y
Soybeans   24 24 25 32 y
Soybeans   21 25 23 24 y
Soybeans   27 45 24 12 y
Soybeans   12 13 15 42 y
Soybeans   22 32 31 43 y
Cotton     31 32 33 34 t
Cotton     29 24 26 28 t
Cotton     34 32 28 45 t
Cotton     26 25 23 24 t
Cotton     53 48 75 26 t
Cotton     34 35 25 78 t
Sugarbeets 22 23 25 42 g
Sugarbeets 25 25 24 26 g
Sugarbeets 34 25 16 52 g
Sugarbeets 54 23 21 54 g
Sugarbeets 25 43 32 15 g
Sugarbeets 26 54  2 54 g
Clover     12 45 32 54 l
Clover     24 58 25 34 l
Clover     87 54 61 21 l
Clover     51 31 31 16 l
Clover     96 48 54 62 l
Clover     31 31 11 11 l
Clover     56 13 13 71 l
Clover     32 13 27 32 l
```

```
   Clover    36 26 54 32 1
   Clover    53 08 06 54 1
   Clover    32 32 62 16 1
```

The SAS code and output:

```
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;

goptions reset=all;
symbol1 c=blue v=triangle;
symbol2 c=cyan v=circle;
symbol3 c=red v=diamond;
symbol4 c=black v=plus;
symbol5 c=green v=x;

proc gplot;
plot Can1 * Can2 = Crop;
run;
```

```
The DISCRIM Procedure
Total Sample Size       36          DF Total                35
Variables                4          DF Within Classes       31
Classes                  5          DF Between Classes       4

Number of Observations Read             36
Number of Observations Used             36
```

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

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

The DISCRIM Procedure
Test of Homogeneity of Within Covariance Matrices
Notation: K = Number of Groups
P = Number of Variables
N = Total Number of Observations - Number of Groups
N(i) = Number of Observations in the i'th Group - 1

$$V = \frac{\prod_{i} |\text{Within SS Matrix}(i)|^{N(i)/2}}{|\text{Pooled SS Matrix}|^{N/2}}$$

$$RHO = 1.0 - \left(\sum_{i} \frac{1}{N(i)} - \frac{1}{N} \right) \frac{2P^2 + 3P - 1}{6(P+1)(K-1)}$$

$$DF = .5(K-1)P(P+1)$$

Under the null hypothesis:

$$-2\ RHO \ln \left(\frac{N^{PN/2}}{\prod_{i} N(i)^{PN(i)/2}} V \right)$$

is distributed approximately as Chi-Square(DF).

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

Since the Chi-Square value is significant at the 0.1 level, the within covariance matrices w
sed in the discriminant function.
Reference: Morrison, D.F. (1976) Multivariate Statistical Methods p252.

The DISCRIM Procedure
Pairwise Generalized Squared Distances Between Groups

$$D^2(i|j) = (\bar{X}_i - \bar{X}_j)' COV_j^{-1} (\bar{X}_i - \bar{X}_j) + \ln |COV_j|$$

Generalized Squared Distance to Crop

From Crop	Clover	Corn	Cotton	Soybeans	Sugarbee
Clover	23.64618	1317	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

The DISCRIM Procedure
Canonical Discriminant Analysis

	Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation
1	0.634584	0.546841	0.100963	0.402697
2	0.392116	0.268638	0.143042	0.153755
3	0.223852	0.147462	0.160561	0.050110
4	0.082467	.	0.167881	0.006801

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

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

	Eigenvalue	Difference	Proportion	Cumulative	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr
1	0.6742	0.4925	0.7364	0.7364	0.47687044	1.48	16	86.179	0.
2	0.1817	0.1289	0.1985	0.9349	0.79837318	0.76	9	70.729	0.
3	0.0528	0.0459	0.0576	0.9925	0.94343017	0.44	4	60	0.
4	0.0068		0.0075	1.0000	0.99319917	0.21	1	31	0.

The DISCRIM Procedure
Canonical Discriminant Analysis

Total Canonical Structure

Variable	Can1	Can2	Can3	Can4
x1	0.965974	0.208737	-0.104594	-0.111282
x2	0.467317	-0.120604	0.408382	0.774787
x3	0.307180	0.829294	0.025572	0.466109
x4	0.200718	-0.063111	0.692647	-0.689905

Between Canonical Structure

Variable	Can1	Can2	Can3	Can4
x1	0.990385	0.132240	-0.037828	-0.014827
x2	0.925735	-0.147626	0.285374	0.199457
x3	0.511475	0.853227	0.015020	0.100858
x4	0.606441	-0.117824	0.738220	-0.270884

Pooled Within Canonical Structure

Variable	Can1	Can2	Can3	Can4
x1	0.950499	0.244477	-0.129788	-0.141199
x2	0.381259	-0.117118	0.420160	0.815102
x3	0.256786	0.825158	0.026957	0.502442
x4	0.158665	-0.059382	0.690471	-0.703241

The DISCRIM Procedure
Canonical Discriminant Analysis

Total-Sample Standardized Canonical Coefficients

Variable	Can1	Can2	Can3	Can4
x1	1.185609267	-0.177733211	-0.576101660	-0.283136415
x2	0.320968184	-0.539432840	0.583201942	0.690577893
x3	-0.251814280	1.218669350	0.302280124	0.137072549
x4	-0.000909970	0.246216548	0.952105109	-0.454977546

Pooled Within-Class Standardized Canonical Coefficients

Variable	Can1	Can2	Can3	Can4
x1	0.989476376	-0.148331173	-0.480798353	-0.236297743
x2	0.323075186	-0.542973959	0.587030384	0.695111207
x3	-0.247373396	1.197177443	0.296949247	0.134655199
x4	-0.000945330	0.255784076	0.989102186	-0.472657148

Raw Canonical Coefficients

Variable	Can1	Can2	Can3	Can4
x1	0.0614736001	-.0092154309	-.0298707542	-.0146805657
x2	0.0254896366	-.0428389722	0.0463148884	0.0548421322
x3	-.0164212569	0.0794715954	0.0197122244	0.0089387447
x4	-.0000514362	0.0139174233	0.0538178684	-.0257176665

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

The DISCRIM Procedure
Classification Results for Calibration Data: WORK.CROPS
Resubstitution Results using Quadratic Discriminant Function
Generalized Squared Distance Function

$$D_j^2(X) = (X-\bar{X}_j)' \, COV_j^{-1} \, (X-\bar{X}_j) + \ln |COV_j|$$

Posterior Probability of Membership in Each Crop

$$Pr(j|X) = \exp(-.5\, D_j^2(X)) \,/\, \sum_k \exp(-.5\, D_k^2(X))$$

Posterior Probability of Membership in Crop

From	Classified

Obs	Crop	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
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

* Misclassified observation

The DISCRIM Procedure
Classification Summary for Calibration Data: WORK.CROPS
Resubstitution Summary using Quadratic Discriminant Function
Generalized Squared Distance Function

$$D_j(X) = (X - \bar{X}_j)' \, COV_j^{-1} \, (X - \bar{X}_j) + \ln |COV_j|$$

Posterior Probability of Membership in Each Crop

$$\Pr(j|X) = \exp(-.5\ D_j(X)) \,/\, \sum_k \exp(-.5\ D_k(X))$$

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
Priors	0.2	0.2	0.2	0.2	0.2	

Error Count Estimates for Crop

	Clover	Corn	Cotton	Soybeans	Sugarbee	Total
Rate	0.1818	0.0000	0.0000	0.0000	0.3333	0.1030
Priors	0.2000	0.2000	0.2000	0.2000	0.2000	

The DISCRIM Procedure
Classification Results for Calibration Data: WORK.CROPS
Cross-validation Results using Quadratic Discriminant Function
Generalized Squared Distance Function

$$D_j(X) = (X-\bar{X}_{(X)j})'\ \mathrm{COV}_{(X)j}^{-1}\ (X-\bar{X}_{(X)j}) + \ln|\mathrm{COV}_{(X)j}|$$

Posterior Probability of Membership in Each Crop

$$\Pr(j|X) = \exp(-.5\ D_j(X)) \,/\, \sum_k \exp(-.5\ D_k(X))$$

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
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

* Misclassified observation

The DISCRIM Procedure
Classification Summary for Calibration Data: WORK.CROPS
Cross-validation Summary using Quadratic Discriminant Function
Generalized Squared Distance Function

$$D_j(X) = (X-\bar{X}_{(X)j})' \, COV_{(X)j}^{-1} \, (X-\bar{X}_{(X)j}) + \ln |COV_{(X)j}|$$

Posterior Probability of Membership in Each Crop

$$Pr(j|X) = \exp(-.5\, D_j(X)) \,/\, SUM_k \exp(-.5\, D_k(X))$$

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
Priors	0.2	0.2	0.2	0.2	0.2	

Error Count Estimates for Crop

	Clover	Corn	Cotton	Soybeans	Sugarbee	Total
Rate	0.1818	0.7143	0.6667	0.6667	0.8333	0.6126
Priors	0.2000	0.2000	0.2000	0.2000	0.2000	