Chapter 10

10.1. $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2} = \begin{bmatrix} 0 & 0 \\ 0 & (.95)^2 \end{bmatrix}$

which has eigenvalues $\rho_1^{*2} = (.95)^2$ and $\rho_2^{*2} = 0$.

The normalized eigenvectors are $\underset{\sim}{e}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\underset{\sim}{e}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Thus

$$U_1 = \underset{\sim}{e}_1' \Sigma_{11}^{1/2} \underset{\sim}{X}^{(1)} = [0 \;\; 1] \begin{bmatrix} .1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_1^{(1)} \\ X_2^{(1)} \end{bmatrix} = X_2^{(1)}$$

Since $\underset{\sim}{f}_1' \Sigma_{22}^{-1/2} = [1 \;\; 0]$, $V_1 = X_1^{(2)}$.

Thus $U_1 = X_2^{(1)}$, $V_1 = X_1^{(2)}$ and $\rho_1^* = .95$.

10.2 a) $\rho_1^* = .55$, $\rho_2^* = .49$

b) $U_1 = .32X_1^{(1)} - .36X_2^{(1)}$

$V_1 = .36X_1^{(2)} - .10X_2^{(2)}$

$U_2 = .20X_1^{(1)} + .30X_2^{(1)}$

$V_2 = .23X_1^{(2)} + .30X_2^{(1)}$

c) $$E\begin{bmatrix} U_1 \\ U_2 \\ V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} -1.675 \\ .015 \\ -.095 \\ .386 \end{bmatrix}$$

$$\text{Cov}\begin{bmatrix} \underset{\sim}{U} \\ \underset{\sim}{V} \end{bmatrix} = \begin{bmatrix} 1 & 0 & .55 & 0 \\ 0 & 1 & 0 & .49 \\ .55 & 0 & 1 & 0 \\ 0 & .49 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \rho_1^* & 0 \\ 0 & 1 & 0 & \rho_2^* \\ \rho_1^* & 0 & 1 & 0 \\ 0 & \rho_2^* & 0 & 1 \end{bmatrix}$$

10.5 a) $$\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} = \rho_{11}^{-1}\rho_{12}\rho_{22}^{-1}\rho_{21} = \begin{bmatrix} .45189 & .28919 \\ .14633 & .17361 \end{bmatrix}$$

$$\begin{vmatrix} .45189-\lambda & .28919 \\ .14633 & .17361-\lambda \end{vmatrix} = \lambda^2 - .5467\lambda + .0005$$
$$= (\lambda - .5457)(\lambda - .0009)$$

The characteristic equation is the same as that of ${}_{11}^{-1/2}\;{}_{12}\;{}_{22}^{-1}\;{}_{21}\;{}_{11}^{-1/2}$ (see Example 10.1) and consequently the eigenvalues are the same.

b) $U_2 = -.677Z_1^{(1)} + 1.055Z_2^{(1)}$

$V_2 = -.863Z_1^{(2)} + .706Z_2^{(2)}$

$\text{Var}(U_2) = (-0.677)^2 + (1.055)^2 - 2(.677)(1.055)(.4) = 1.0$

$\text{Var}(V_2) = 1.0$

$\text{Corr}(U_2, V_2) = (-.677)(-.863)(.5) + (-.863)(1.055)(.3)$
$+ (.706)(-.677)(.6) + (.706)(1.055)(.4) = .03 = \rho_2^*$

10.7 a) $\rho_1^* = \frac{2\rho}{1+\rho} \quad 0 < \rho < 1$

$$U_1 = \frac{1}{\sqrt{2(1+\rho)}} (X_1^{(1)}+X_2^{(1)})$$

$$V_1 = \frac{1}{\sqrt{2(1+\rho)}} (X_1^{(2)}+X_2^{(2)})$$

10.8 c) $\hat{\rho}_1^* = .72$

$\hat{V}_1 = .20X_1^{(2)}+.70X_2^{(2)}$

$\hat{\beta} = 45^\circ \equiv \frac{\pi}{4}$ radians

d) $\hat{\rho}_1^* = .57$

$\hat{U}_1 = 1.03 \cos \theta_1 + .46 \sin \theta_1$

$V_1 = .49 \cos \theta_2 + .78 \sin \theta_2$

10.9 a) $\hat{\rho}_1^* = .39 \quad ; \quad \hat{\rho}_2^* = .07$

$\hat{U}_1 = 1.26Z_1^{(1)}-1.03Z_2^{(1)}; \quad \hat{U}_2 = .30Z_1^{(1)}+.79Z_2^{(1)}$

$\hat{V}_1 = 1.10Z_1^{(2)}-.45Z_2^{(2)} \; ; \quad V_2 = -.02Z_1^{(2)}+1.01Z_2^{(2)}$

b) $n = 140$, $p=2$, $q=2$, $n-1-\frac{1}{2}(p+q+1) = 136.5$

Null hypothesis	Value of test statistic	Degrees of Freedom	Upper 5% point of χ^2 distribution
$H_0: \Sigma_{12} = \rho_{12} = 0$	$-136.5 \ln(.8444)(.9953) = 23.74$	4	9.48
$H_0^{(1)}: \rho_1^* \neq 0,\ \rho_2^* = 0$	$-136.5 \ln(.9953) = .65$	1	3.84

Therefore, reject H_0 but do not reject $H_0^{(1)}$. Reading ability (summarized by $\hat{U}_1$) does correlate with arithmetic ability (summarized by $\hat{V}_1$) but the correlation (represented by $\rho_1 = .39$) is not particularly strong.

10.10 a) $\hat{\rho}_1^* = .33, \ \hat{\rho}_2^* = .17$

b) $\hat{U}_1 = 1.002Z_1^{(1)} - .003Z_2^{(1)}$

$\hat{V}_1 = -.602Z_1^{(2)} - .977Z_2^{(2)}$

$\hat{U}_1 \doteq Z_1^{(1)}$ = 1973 nonprimary homicides (standardized)

$\hat{V}_1 \doteq \frac{3}{5} Z_1^{(2)} + Z_2^{(2)}$ = a "punishment index"

Punishment appears to be correlated with nonprimary homicides but not primary homicides.

10.11 Using the correlation matrix **R** and standardized variables, the canonical correlations and canonical variables follow. The $Z^{(1)}$'s are the banks, the $Z^{(2)}$'s are the oil companies.

$\hat{\rho}_1^* = .348, \quad \hat{\rho}_2^* = .130$

$$\hat{U}_1 = -.539z_1^{(1)} + 1.209z_2^{(1)} + .079z_3^{(1)}, \qquad \hat{U}_2 = 1.142z_1^{(1)} - .410z_2^{(1)} + .142z_3^{(1)}$$
$$\hat{V}_1 = 1.160z_1^{(2)} - .261z_2^{(2)} \qquad\qquad \hat{V}_2 = -.728z_1^{(2)} + 1.345z_2^{(2)}$$

Additional correlations:

$$\mathbf{R}_{U_1,z^{(1)}} = (.266 \ \ .913 \ \ .498), \quad \mathbf{R}_{V_1,z^{(2)}} = (.982 \ \ .532)$$
$$\mathbf{R}_{U_1,z^{(2)}} = (.342 \ \ .185), \quad \mathbf{R}_{V_1,z^{(1)}} = (.093 \ \ .318 \ \ .174)$$

Here $H_0 : \boldsymbol{\Sigma}_{12}\,(\boldsymbol{\rho}_{12}) = \mathbf{0}$ is rejected at the 5% level and $H_0^{(1)} : \rho_1^* \neq 0, \rho_2^* = 0$ is not rejected at the 5% level. The first canonical correlation, although relatively small, is significant. The second canonical correlation is not significant.

Focusing attention on the first pair of canonical variables, $\hat{U}_1$ is dominated by Citibank, $\hat{V}_1$ is dominated by Royal Dutch Shell. The canonical correlation (.348) between $\hat{U}_1$ and $\hat{V}_1$ suggests there is not much co-movement between the rates of return for the banks on one hand and the oil companies on the other. Moreover, $\hat{U}_1$ is not highly correlated with any of the $Z^{(2)}$'s (oil companies) and $\hat{V}_1$ is not highly correlated with any of the $Z^{(1)}$'s (banks). The first canonical variables differentiate stocks in different industries with some, but not much, overlap.

10.12 a) $\hat{\rho}_1^* = .69$, $\hat{\rho}_2^* = .19$

Reject H_0: $\rho_{12} = 0$ at the 5% level but do not reject $H_0^{(1)} = \rho_1^* \neq 0$, $\rho_2^* = 0$ at the 5% level.

b) $\hat{U}_1 = .77Z_1^{(1)} + .27Z_2^{(1)}$

$\hat{V}_1 = .05Z_1^{(2)} + .90Z_2^{(2)} + .19Z_3^{(2)}$

c) Sample Correlations Between Original Variables and Canonical Variables

$\underset{\sim}{X}^{(1)}$ Variables	$\hat{U}_1$	$\hat{V}_1$	$\underset{\sim}{X}^{(2)}$ Variables	$\hat{U}_1$	$\hat{V}_1$
1. annual frequency of restaurant dining	.99	.68	1. age of head of household	.29	.42
2. annual frequency of attending movies	.89	.61	2. annual family income	.68	.98
			3. educational level of head of household	.35	.51

d) $\hat{U}_1$ is a measure of family entertainment outside the home. $\hat{V}_1$ may be considered a measure of family "status" which is dominated by family income. Essentially, family entertainment outside the home is positively associated with family income.

10.13 a) $\hat{\rho}_1^* = .909$, $\hat{\rho}_2^* = .636$, $\hat{\rho}_3^* = .256$, $\hat{\rho}_4^* = .094$

	Null hypothesis	Value of test statistic	Degrees of freedom	Conclusion at 1% level
1.	H_0: $\Sigma_{12} = \rho_{12} = 0$	309.98	20	Reject H_0
2	H_0: $\rho_1 \neq 0$, $\rho_2 = \ldots = \rho_4 = 0$	78.63	12	Reject H_0
3.	H_0: $\rho_1 \neq 0$, $\rho_2 \neq 0$, $\rho_3 = 0$, $\rho_4 = 0$	16.81	6	Do not reject H_0.

$$\begin{bmatrix} \hat{U}_1 \\ \hat{U}_2 \end{bmatrix} = \begin{bmatrix} .21 & .17 & -.33 & -.26 & .30 \\ .92 & -.58 & .65 & .34 & .55 \end{bmatrix} \begin{bmatrix} Z_1^{(1)} \\ Z_2^{(1)} \\ Z_3^{(1)} \\ Z_4^{(1)} \\ Z_5^{(1)} \end{bmatrix}$$

$$\begin{bmatrix} \hat{V}_1 \\ \hat{V}_2 \end{bmatrix} = \begin{bmatrix} -.54 & -.29 & .46 & .03 \\ 1.01 & .03 & .98 & -.18 \end{bmatrix} \begin{bmatrix} Z_1^{(2)} \\ Z_2^{(2)} \\ Z_3^{(2)} \\ Z_4^{(2)} \end{bmatrix}$$

b) $\hat{U}_1$ appears to measure quality of wheat as a "contrast" between "good" aspects ($Z_1^{(1)}$, $Z_2^{(1)}$ and $Z_5^{(1)}$) and "bad" aspects ($Z_3^{(1)}$, $Z_4^{(1)}$).

$\hat{V}_1$ is harder to interpret. It appears to measure the quality of the flour as represented by $Z_1^{(2)}$, $Z_2^{(2)}$ and $Z_3^{(2)}$.

10.14

a) $\hat{\rho}_1^* = 0.7520$, $\hat{\rho}_2^* = 0.5395$. And the sample canonical variates are

Raw Canonical Coefficients for the Accounting measures of profitability

	U1	U2
HRA	-0.494697741	1.9655018549
HRE	0.2133051339	-0.794353012
HRS	0.7228315515	-0.538822808
RRA	2.7749354333	-4.38345956
RRE	-1.383659039	1.6471230054
RRS	-1.032933813	2.5190103052

Raw Canonical Coefficients for the Market measures of profitability

	V1	V2
Q	1.3930601511	-2.500804367
REV	-0.431692979	2.8298904995

U_1 is most highly correlated with RRA and HRA and also HRS and RRS. V_1 is highly correlated with both of its components. The second pair does not correlate well with their respective components.

b) Standardized Variance of the Accounting measures of profitability Explained by

	Their Own Canonical Variables			The Opposite Canonical Variables	
	Proportion	Cumulative Proportion	Canonical R-Squared	Proportion	Cumulative Proportion
1	0.5041	0.5041	0.5655	0.2851	0.2851
2	0.0905	0.5946	0.2910	0.0263	0.3114

Standardized Variance of the Market measures of profitability Explained by

	Their Own Canonical Variables			The Opposite Canonical Variables	
	Proportion	Cumulative Proportion	Canonical R-Squared	Proportion	Cumulative Proportion
1	0.8702	0.8702	0.5655	0.4921	0.4921
2	0.1298	1.0000	0.2910	0.0378	0.5299

Market measures can be well explained by its canonical variate $\hat{V}_1$. However, accounting measures cannot be well explained. In fact, from the correlation between measures and canonical variates, accounting measures on equity have weak correlation with $\hat{U}_1$.

Correlations Between the Accounting measures of profitability and Their Canonical Variables

	U1	U2
HRA	0.8110	0.2711
HRE	0.4225	0.0968
HRS	0.7184	0.5526
RRA	0.8605	-0.0089

RRE	0.5741	-0.0959
RRS	0.7761	0.3814

Correlations Between the Market measures of profitability and Their Canonical Variables

	V1	V2
Q	0.9886	0.1508
REV	0.8736	0.4866

10.15

$\hat{\rho}_1^* = 0.9129$, $\hat{\rho}_2^* = 0.0681$. And the sample canonical variates are

Raw Canonical Coefficients for the dynamic measure

	U1	U2
X1	0.0036015621	-0.006663216
X2	-0.000595735	0.0077029513

Raw Canonical Coefficients for the static measures

	V1	V2
X3	0.0013448038	0.008471035
X4	0.0018933921	-0.007828962

Standardized Variance of the dynamic measure Explained by

	Their Own Canonical Variables			The Opposite Canonical Variables	
	Proportion	Cumulative Proportion	Canonical R-Squared	Proportion	Cumulative Proportion
1	0.8840	0.8840	0.8334	0.7367	0.7367
2	0.1160	1.0000	0.0046	0.0005	0.7373

Standardized Variance of the static measures Explained by

	Their Own Canonical Variables			The Opposite Canonical Variables	
	Proportion	Cumulative Proportion	Canonical R-Squared	Proportion	Cumulative Proportion
1	0.9601	0.9601	0.8334	0.8002	0.8002
2	0.0399	1.0000	0.0046	0.0002	0.8003

Static measures can be well explained by its canonical variate $\hat{U}_1$. Also, dynamic measures can be well explained by its canonical variate $\hat{V}_1$.

10.16 From the computer output below, the first two canonical correlations are $\hat{\rho}_1^* = 0.517345$ and $\hat{\rho}_2^* = 0.125508$. The large sample tests

$$-(n-1-\frac{1}{2}(p+q-1)\ )\ln[(1-\hat{\rho}*_1^2)(1-\hat{\rho}*_1^2)\ ]\ \geq\ \chi^2_{pq}(.05)$$

or

$$-(46-1-\frac{1}{2}(3+2-1)\ )\ln[(1-(.517345)^2)(1-(.125508)^2)\ ]\ =\ 13.50\ \geq\ \chi^2_6(.05)=12.59$$

and

$$-(n-1-\frac{1}{2}(p+q-1)\ )\ln[(1-\hat{\rho}*_1^2)\ ]\ \geq\ \chi^2_{(p-1)(q-1)}(.05)$$

or

$$-(46-1-\frac{1}{2}(3+2-1)\ )\ln[(1-(.125508)^2)\ ]\ =\ 0.667\ \geq\ \chi^2_2(.05)=5.99$$

suggest that only the first pair of canonical variables are important. Even if the variables means were given, we prefer to interpret the canonical variables obtained from $\mathbf{S}$ in terms of coefficients of standardized variables.

$$\begin{aligned}\hat{U}_1 &= .4357z_1^{(1)} - .7047z_2^{(1)} + 1.0815z_3^{(1)}\\ \hat{V}_1 &= 1.020z_1^{(2)} - .1609z_2^{(2)}\end{aligned}$$

The two insulin responses dominate $\hat{U}_1$ while $\hat{V}_1$ consists primarily of the relative weight variable.

Canonical Correlation Analysis

	Canonical Correlation	Adjusted Canonical Correlation	Approx Standard Error	Squared Canonical Correlation
1	0.517345	0.517145	0.007324	0.267646
2	0.125508	0.125158	0.009843	0.015752

Canonical Correlation Analysis

Raw Canonical Coefficients for the Glucose and Insulin

GLUCOSE	0.0131006541	0.0247524811
INSULIN	-0.014438254	-0.009317525
INSULRES	0.023399723	-0.008667216

Raw Canonical Coefficients for the Weight and Fasting

WEIGHT	8.0655750801	-0.375167814
FASTING	-0.019159052	0.1200675138

Standardized Canonical Coefficients for the Glucose and Insulin

GLUCOSE	0.4357	0.8232
INSULIN	-0.7047	-0.4547
INSULRES	1.0815	-0.4006

Standardized Canonical Coefficients for the Weight and Fasting

	SECONDA1	SECONDA2
WEIGHT	1.0202	-0.0475
FASTING	-0.1609	1.0086

Correlations Between the Glucose and Insulin and Their Canonical Variables

	PRIMARY1	PRIMARY2
GLUCOSE	0.3397	0.6838
INSULIN	-0.0502	-0.4565
INSULRES	0.7551	-0.5729

Correlations Between the Weight and Fasting and Their Canonical Variables

	SECONDA1	SECONDA2
WEIGHT	0.9875	0.1576
FASTING	0.0465	0.9989

10.17 The computer output below suggests maybe two canonical pairs of variables. the canonical correlations are 0.521594, 0.375256, 0.242181 and 0.136568. $\hat{U}_1$ ignores the first smoking question and $\hat{U}_2$ ignores the third. $\hat{V}_1$ depends heavily on the difference of annoyance and tenseness.

Even the second pairs do not explain their own variances very well. $R^2_{\boldsymbol{z}^{(1)}|\hat{U}_2} = .1249$ and $R^2_{\boldsymbol{z}^{(1)}|\hat{V}_2} = 0.0879$

Canonical Correlation Analysis

	Canonical Correlation	Adjusted Canonical Correlation	Approx Standard Error	Squared Canonical Correlation
1	0.521594	0.520771	0.007280	0.272060
2	0.375256	0.374364	0.008592	0.140817
3	0.242181	0.241172	0.009414	0.058652
4	0.136568	0.135586	0.009814	0.018651

Standardized Canonical Coefficients for the Smoking

	SMOKING1	SMOKING2	SMOKING3	SMOKING4
SMOK1	-0.0430	1.0898	1.1161	-1.0092
SMOK2	1.1622	0.6988	-1.4170	0.1732
SMOK3	-1.3753	0.2081	0.0156	1.6899
SMOK4	0.8909	-1.6506	0.8325	-0.2630

Standardized Canonical Coefficients for the Psych and Physical State

	STATE1	STATE2	STATE3	STATE4
CONCEN	0.4733	-0.8141	0.4946	-0.1604
ANNOY	-0.7806	-0.4510	0.5909	-0.7193
SLEEP	0.2567	-0.6052	0.6981	0.6246
TENSE	0.6919	0.3800	-0.4190	0.4376
ALERT	-0.1451	-0.1840	-1.5191	-0.7253
IRRITAB	-0.0704	0.6255	-0.3343	0.8760
TIRED	0.3127	0.5898	0.2276	0.1861
CONTENT	0.3364	0.4869	0.8334	-0.6557

Canonical Structure

Correlations Between the Smoking and Their Canonical Variables

	SMOKING1	SMOKING2	SMOKING3	SMOKING4
SMOK1	0.4458	0.5278	0.6615	0.2917
SMOK2	0.7305	0.3822	0.1487	0.5461
SMOK3	0.2910	0.2664	0.4668	0.7915
SMOK4	0.6403	-0.0620	0.5586	0.5236

Correlations Between the Psychological and Physical State and Their Canonical Variables

	STATE1	STATE2	STATE3	STATE4
CONCEN	0.7199	-0.3579	0.0125	-0.3137
ANNOY	0.3035	0.1365	0.3906	-0.4058
SLEEP	0.5995	-0.3490	0.3709	0.2586
TENSE	0.7015	0.3305	0.0053	-0.1861
ALERT	0.7290	-0.1539	-0.1459	-0.3681
IRRITAB	0.4585	0.3342	0.1211	-0.0805
TIRED	0.6905	-0.0267	0.2544	0.0749
CONTENT	0.5323	0.4350	0.3207	-0.5601

Canonical Redundancy Analysis
Raw Variance of the Smoking
Explained by

	Their Own Canonical Variables			The Opposite Canonical Variables	
	Proportion	Cumulative Proportion	Canonical R-Squared	Proportion	Cumulative Proportion
1	0.3068	0.3068	0.2721	0.0835	0.0835
2	0.1249	0.4316	0.1408	0.0176	0.1010
3	0.2474	0.6790	0.0587	0.0145	0.1155
4	0.3210	1.0000	0.0187	0.0060	0.1215

Raw Variance of the Psychological and Physical State
Explained by

	Their Own Canonical Variables			The Opposite Canonical Variables	
	Proportion	Cumulative Proportion	Canonical R-Squared	Proportion	Cumulative Proportion
1	0.3705	0.3705	0.2721	0.1008	0.1008
2	0.0879	0.4583	0.1408	0.0124	0.1132
3	0.0617	0.5201	0.0587	0.0036	0.1168
4	0.1032	0.6233	0.0187	0.0019	0.1187

10.18 The canonical correlation analysis expressed in terms of standardized variables follows. The $Z^{(1)}$'s are the paper characteristic variables, the $Z^{(2)}$'s are the pulp fiber characteristic variables.

Canonical correlations:

$$\hat{\rho}_1^* = .917, \quad \hat{\rho}_2^* = .817, \quad \hat{\rho}_3^* = .265, \quad \hat{\rho}_4^* = .092$$

First three canonical variate pairs:

$$\hat{U}_1 = -1.505z_1^{(1)} - .212z_2^{(1)} + 1.998z_3^{(1)} + .676z_4^{(1)}$$
$$\hat{V}_1 = -.159z_1^{(2)} + .633z_2^{(2)} + .325z_3^{(2)} + .818z_4^{(2)}$$

$$\hat{U}_2 = -3.496z_1^{(1)} - 1.543z_2^{(1)} + 1.076z_3^{(1)} + 3.768z_4^{(1)}$$
$$\hat{V}_2 = .689z_1^{(2)} + 1.003z_2^{(2)} + .005z_3^{(2)} - 1.562z_4^{(2)}$$

$$\hat{U}_3 = -5.702z_1^{(1)} + 3.525z_2^{(1)} - 4.714z_3^{(1)} + 7.153z_4^{(1)}$$
$$\hat{V}_3 = -.513z_1^{(2)} + .077z_2^{(2)} - 1.663z_3^{(2)} - .779z_4^{(2)}$$

Additional correlations:

$$R_{U_1,Z^{(1)}} = (.935\ \ .887\ \ .977\ \ .952),\quad R_{V_1,Z^{(2)}} = (.817\ \ .906\ \ -.650\ \ .940)$$

$$R_{U_1,Z^{(2)}} = (.749\ \ .831\ \ -.596\ \ .862),\quad R_{V_1,Z^{(1)}} = (.858\ \ .814\ \ .896\ \ .873)$$

Here $H_0:\Sigma_{12}\,(\rho_{12}) = \mathbf{0}$ is rejected at the 5% level and $H_0^{(1)}:\rho_1^* \neq 0, \rho_2^* = 0$ is rejected at the 5% level. $H_0^{(2)}:\rho_1^* \neq 0, \rho_2^* \neq 0, \rho_3^* = \rho_4^* = 0$ is not rejected at the 5% level. The first two canonical correlations are significantly different from 0. The last two canonical correlations are not significant.

The first canonical variable $\hat{U}_1$ explains 88% of the total standardized variance of it's set, the $Z^{(1)}$'s. The first canonical variable $\hat{V}_1$ explains 70% of the total standardized variance of it's set, the $Z^{(2)}$'s. The first canonical variates are good summary measures of their respective sets of variables. Moreover, the first canonical variates, which might be labeled a "paper characteristic index" and "a pulp fiber strength—quality index", are highly correlated. There is a strong association between an index of pulp fiber characteristics and an index of the characteristics of paper made from them.

The second canonical variable $\hat{U}_2$ appears to be a contrast between the first two variables, breaking length and elastic modulus, and the last two variables, stress at failure and burst strength. However, the only moderately large (in absolute value) correlation between the canonical variate and it's component variables is the correlation (−.428) between $\hat{U}_2$ and $Z_2^{(1)}$, elastic modulus. The remaining correlations are small. This canonical variable might be a "paper stretch" measure. The canonical variable $\hat{V}_2$ appears to be determined by all variables except $Z_3^{(2)}$, fine fiber fraction. This canonical variable might be a "fiber length/strength" measure. The second pair of canonical variates is also highly correlated.

10.19 The correlation matrix **R** and the canonical analysis for the standardized variables follows. The $Z^{(1)}$'s are the running speed events (100m, 400m, long jump), the $Z^{(2)}$'s are the arm strength events (discus, javelin, shot put).

$$\mathbf{R}_{11} = \begin{pmatrix} 1.0 & .5520 & .6386 \\ .5520 & 1.0 & .4706 \\ .6386 & .4706 & 1.0 \end{pmatrix} \qquad \mathbf{R}_{22} = \begin{pmatrix} 1.0 & .4179 & .7926 \\ .4179 & 1.0 & .4682 \\ .7926 & .4682 & 1.0 \end{pmatrix}$$

$$\mathbf{R}_{12} = \mathbf{R}_{21}' = \begin{pmatrix} .3509 & .1821 & .4752 \\ .2100 & .2116 & .2539 \\ .3998 & .3102 & .4953 \end{pmatrix}$$

Canonical correlations:

$$\hat{\rho}_1^* = .540, \quad \hat{\rho}_2^* = .212, \quad \hat{\rho}_3^* = .014$$

Canonical variables:

$$\hat{U}_1 = .540z_1^{(1)} - .120z_2^{(1)} + .633z_3^{(1)} \qquad \hat{U}_2 = 1.277z_1^{(1)} - .768z_2^{(1)} - .773z_3^{(1)}$$

$$\hat{V}_1 = -.057z_1^{(2)} + .043z_2^{(2)} + 1.024z_3^{(2)} \qquad \hat{V}_2 = -.422z_1^{(2)} - 1.0685z_2^{(2)} + .859z_3^{(2)}$$

$$\hat{U}_3 = .399z_1^{(1)} + .940z_2^{(1)} - .866z_3^{(1)}$$

$$\hat{V}_3 = 1.590z_1^{(2)} - .384z_2^{(2)} - 1.038z_3^{(2)}$$

Additional correlations:

$$R_{U_1,z^{(1)}} = (.662 \;\; .160 \;\; .732), \quad R_{V_1,z^{(2)}} = (.772 \;\; .498 \;\; .999)$$

Here $H_0 : \boldsymbol{\Sigma}_{12}\,(\boldsymbol{\rho}_{12}) = \mathbf{0}$ is rejected at the 5% level and $H_0^{(1)} : \rho_1^* \neq 0, \rho_2^* = \rho_3^* = 0$ is rejected at the 5% level. $H_0^{(2)} : \rho_1^* \neq 0, \rho_2^* \neq 0, \rho_3^* = 0$ is not rejected at the 5% level. The first and second canonical correlations are significant. The third canonical correlation is not significant.

We might identify $\hat{U}_1$ as a "running speed" measure since the 100m run and the long jump receive the greatest weight in this canonical variate and also are each highly correlated with $\hat{U}_1$. We might call $\hat{V}_1$ a "strength" or "arm strength" measure since the shot put has a large coefficient in this canonical variate and the discuss, javelin and shot put are each highly correlated with $\hat{V}_1$.