NFL Combine Example – Canonical Correlation Analysis

$\mathbf{X}^{(1)} \equiv$ Body Dimension Variables: Height, Arm Length, Weight, Hand Length

$\mathbf{X}^{(2)} \equiv$ Physical Performance Variables: 40 Yard Time, 225 lb Bench Press Reps, Vertical Jump, Broad Jump

Correlation Matrix: $\mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix}$

```
        height  armlen      wt handlen  time40   bench   vjump   bjump
height   1.0000  0.7644  0.7208  0.4064  0.5895  0.3871 -0.4425 -0.4218
armlen   0.7644  1.0000  0.6347  0.4766  0.4812  0.3333 -0.3828 -0.3514
wt       0.7208  0.6347  1.0000  0.4619  0.8910  0.7123 -0.7103 -0.7512
handlen  0.4064  0.4766  0.4619  1.0000  0.3874  0.3042 -0.2539 -0.2726
time40   0.5895  0.4812  0.8910  0.3874  1.0000  0.5671 -0.7691 -0.8456
bench    0.3871  0.3333  0.7123  0.3042  0.5671  1.0000 -0.4098 -0.4427
vjump   -0.4425 -0.3828 -0.7103 -0.2539 -0.7691 -0.4098  1.0000  0.8432
bjump   -0.4218 -0.3514 -0.7512 -0.2726 -0.8456 -0.4427  0.8432  1.0000
```

Matrices used to obtain Canonical Variates and (Squared) Correlations:

$\mathbf{R}_{11}^{-1/2}\mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21}\mathbf{R}_{11}^{-1/2}$ and $\mathbf{R}_{22}^{-1/2}\mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{R}_{12}\mathbf{R}_{22}^{-1/2}$

```
> CC.mat1
           [,1]       [,2]      [,3]       [,4]
[1,] 0.10717575 0.05303344 0.1853638 0.05396137
[2,] 0.05303344 0.03296692 0.1166684 0.02490461
[3,] 0.18536382 0.11666837 0.7876753 0.14152964
[4,] 0.05396137 0.02490461 0.1415296 0.03989531

> CC.mat2
           [,1]       [,2]       [,3]       [,4]
[1,]  0.4191117  0.2893379 -0.2097069 -0.2010653
[2,]  0.2893379  0.2554468 -0.1607914 -0.2011895
[3,] -0.2097069 -0.1607914  0.1225738  0.1211551
[4,] -0.2010653 -0.2011895  0.1211551  0.1705809
```

Eigenvalues and Eigenvectors of $\mathbf{R}_{11}^{-1/2}\mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21}\mathbf{R}_{11}^{-1/2}$ $\left(\hat{\rho}_k^2, \hat{\mathbf{e}}_k\right)$ and Eigenvectors of $\mathbf{R}_{22}^{-1/2}\mathbf{R}_{21}\mathbf{R}_{11}^{-1}\mathbf{R}_{12}\mathbf{R}_{22}^{-1/2}$ $\left(\hat{\mathbf{f}}_k\right)$:

```
> eigen(CC.mat1)$val
[1] 0.881941333 0.073108516 0.011072002 0.001591423
> eigen(CC.mat1)$vec
           [,1]       [,2]         [,3]       [,4]
[1,] -0.2476794  0.8492400 -0.006944013 -0.4662597
[2,] -0.1499372  0.3342368 -0.615766637  0.6975930
[3,] -0.9404034 -0.3241813 -0.049144939 -0.0901814
[4,] -0.1783682  0.2489651  0.786363786  0.5365009
> eigen(CC.mat2)$vec
           [,1]        [,2]         [,3]       [,4]
[1,]  0.6626096  0.65062063  0.277057585 -0.2467397
[2,]  0.5251578 -0.38070053  0.271325808  0.7110968
[3,] -0.3574770 -0.07740005  0.921708783 -0.1291213
[4,] -0.3966964  0.65250988 -0.008620857  0.6455916
```

$$\hat{\rho}_1^* = \sqrt{.88194} = .93912 \quad \hat{\rho}_2^* = \sqrt{.07311} = .27039 \quad \hat{\rho}_3^* = \sqrt{.01107} = .10522 \quad \hat{\rho}_4^* = \sqrt{.00159} = .03989$$

$$\hat{\mathbf{a}}_k = \mathbf{R}_{11}^{-1/2}{}'\hat{\mathbf{e}}_k \qquad \hat{\mathbf{b}}_k = \mathbf{R}_{22}^{-1/2}{}'\hat{\mathbf{f}}_k \qquad k = 1,2,3,4$$

```
> round(AB.out,3)
         a1     b1     a2     b2     a3     b3     a4     b4
[1,]  0.115  0.654  1.321  1.809  0.322  0.744 -1.115 -0.279
[2,]  0.134  0.358  0.050 -0.651 -1.069  0.265  1.220  0.929
[3,] -1.156 -0.076 -0.953 -0.392 -0.084  1.701 -0.152 -0.738
[4,]  0.004 -0.047  0.244  1.847  1.027 -0.467  0.507  1.244
```

Canonical Variates:

$$\hat{U}_k = \hat{\mathbf{a}}_k{}'\mathbf{z}^{(1)} \quad \hat{V}_k = \hat{\mathbf{b}}_k{}'\mathbf{z}^{(2)} \quad k = 1,2,3,4 \quad V\left\{\hat{U}_k\right\} = V\left\{\hat{V}_k\right\} = 1 \quad \mathrm{COV}\left\{\hat{U}_k, \hat{U}_l\right\} = \mathrm{COV}\left\{\hat{V}_k, \hat{V}_l\right\} = \mathrm{COV}\left\{\hat{U}_k, \hat{V}_l\right\} = 0$$

$$\hat{\mathbf{A}} = \begin{bmatrix} \hat{\mathbf{a}}_1{}' \\ \vdots \\ \hat{\mathbf{a}}_p{}' \end{bmatrix} \quad \hat{\mathbf{B}} = \begin{bmatrix} \hat{\mathbf{b}}_1{}' \\ \vdots \\ \hat{\mathbf{b}}_q{}' \end{bmatrix} \quad \hat{\mathbf{U}} = \hat{\mathbf{A}}\mathbf{z}^{(1)} \quad \hat{\mathbf{V}} = \hat{\mathbf{B}}\mathbf{z}^{(2)} \quad \hat{\mathbf{A}}^{-1} = \begin{bmatrix} \hat{\mathbf{a}}^{(1)} & \cdots & \hat{\mathbf{a}}^{(p)} \end{bmatrix} \quad \hat{\mathbf{B}}^{-1} = \begin{bmatrix} \hat{\mathbf{b}}^{(1)} & \cdots & \hat{\mathbf{b}}^{(q)} \end{bmatrix}$$

$$\hat{\mathrm{COV}}\left\{\mathbf{z}^{(1)}, \hat{\mathbf{U}}\right\} = \hat{\mathrm{COV}}\left\{\hat{\mathbf{A}}^{-1}\hat{\mathbf{U}}, \hat{\mathbf{U}}\right\} = \hat{\mathbf{A}}^{-1} \qquad \hat{\mathrm{COV}}\left\{\mathbf{z}^{(2)}, \hat{\mathbf{V}}\right\} = \hat{\mathrm{COV}}\left\{\hat{\mathbf{B}}^{-1}\hat{\mathbf{V}}, \hat{\mathbf{V}}\right\} = \hat{\mathbf{B}}^{-1}$$

```
> round(AB.inv.out,4)
      INV(A)                              INV(B)
[1,] -0.6147 0.7717 -0.1389 -0.0862  0.9546  0.1792 -0.0192 -0.2372
[2,] -0.5105 0.5715 -0.3873  0.5126  0.7807 -0.2822  0.1961  0.5219
[3,] -0.9866 0.1439 -0.0566  0.0526 -0.7646  0.0410  0.6265  0.1453
[4,] -0.4198 0.3650  0.6094  0.5650 -0.8220  0.2751  0.2210  0.4471
```

Proportions of Sample Variances of $\mathbf{Z}^{(1)}$ and $\mathbf{Z}^{(2)}$ explained by first r canonical variates.

$$R^2_{z^{(1)}\bullet\hat{U}_1,\ldots\hat{U}_r} = \sum_{i=1}^{r}\left\{\left(\hat{\mathbf{A}}^{-1}\right)'\hat{\mathbf{A}}^{-1}\right\}_{i,i} \qquad R^2_{z^{(2)}\bullet\hat{V}_1,\ldots\hat{V}_r} = \sum_{i=1}^{r}\left\{\left(\hat{\mathbf{B}}^{-1}\right)'\hat{\mathbf{B}}^{-1}\right\}_{i,i}$$

```
> round(AiAiBiBi.out,4)
     INV(A)'INV(A)                          INV(B)'INV(B)
[1,]        1.7879 -1.0613  0.0831 -0.4977        2.7810 -0.3068 -0.5259 -0.2975
[2,]       -1.0613  1.0761 -0.1142  0.4403       -0.3068  0.1891  0.0277 -0.0608
[3,]        0.0831 -0.1142  0.5438  0.1548       -0.5259  0.0277  0.4802  0.2967
[4,]       -0.4977  0.4403  0.1548  0.5922       -0.2975 -0.0608  0.2967  0.5496
```

```
> round(prop.out,4)
     Z1,r=1   r=2    r=3  r=4 Z2,r=1    r=2    r=3 r=4
[1,]  0.447 0.716 0.8519    1 0.6953 0.7425 0.8626   1
```

Test whether $\mathbf{\Sigma_{12} = 0}$ $\quad TS: \quad -\left(n-1-\frac{p+q+1}{2}\right)\ln\left(\prod_{i=1}^{p}\hat{\rho}_i^{*2}\right)$

```
> round(test.out,4)
    Test Stat DF X2(.05) Pr>TS
[1,]  432.8051 16 26.2962     0
```

Testing whether subsequent (smaller) correlations are 0. $H_0 : \rho_{k+1} = ... = \rho_p = 0 \quad k+1 = 1, ..., p$

```
> round(test.out,4)
     k+1 Test Stat DF X2(.05)  Pr>TS
[1,]   1  432.8051 16 26.2962 0.0000
[2,]   2   17.2415  9 16.9190 0.0451
[3,]   3    2.4753  4  9.4877 0.6491
[4,]   4    0.3098  1  3.8415 0.5778
```