8.17. The data on national track records for women are listed in Table 1.7.

(a) Obtain the sample correlation matrix **R** for these data and determine its eigenvalues and eigenvectors.

(b) Determine the first two principal components for the standardized variables. Prepare a table showing the correlations of the standardized variables with the components and the cumulative percentage of total (standardized) sample variance explained by the two components.

(c) Interpret the two principal components obtained in Part b. (Note the first component is essentially a normalized unit vector and might measure the athletic excellence of a given nation. The second component might measure relative strength of a nation at the various running distances.)

(d) Rank the nations based on their score on the first principal component. Does this ranking correspond with your inituitive notion of athletic excellence for the various countries?

8.18. The data on national track records for men are listed in Table 8.3. (See also the national track records for men data on the disc.) Repeat the principal component analysis outlined in Exercise 8.17 for the men. Are the results consistent with those obtained from the women's data?

TABLE 8.3 NATIONAL TRACK RECORDS FOR MEN

Country	100 m (s)	200 m (s)	400 m (s)	800 m (min)	1500 m (min)	5000 m (min)	10,000 m (min)	Marathon (mins)
Argentina	10.39	20.81	46.84	1.81	3.70	14.04	29.36	137.72
Australia	10.31	20.06	44.84	1.74	3.57	13.28	27.66	128.30
Austria	10.44	20.81	46.82	1.79	3.60	13.26	27.72	135.90
Belgium	10.34	20.68	45.04	1.73	3.60	13.22	27.45	129.95
Bermuda	10.28	20.58	45.91	1.80	3.75	14.68	30.55	146.62
Brazil	10.22	20.43	45.21	1.73	3.66	13.62	28.62	133.13
Burma	10.64	21.52	48.30	1.80	3.85	14.45	30.28	139.95
Canada	10.17	20.22	45.68	1.76	3.63	13.55	28.09	130.15
Chile	10.34	20.80	46.20	1.79	3.71	13.61	29.30	134.03
China	10.51	21.04	47.30	1.81	3.73	13.90	29.13	133.53
Colombia	10.43	21.05	46.10	1.82	3.74	13.49	27.88	131.35
Cook Islands	12.18	23.20	52.94	2.02	4.24	16.70	35.38	164.70
Costa Rica	10.94	21.90	48.66	1.87	3.84	14.03	28.81	136.58
Czechoslovakia	10.35	20.65	45.64	1.76	3.58	13.42	28.19	134.32
Denmark	10.56	20.52	45.89	1.78	3.61	13.50	28.11	130.78
Dominican Republic	10.14	20.65	46.80	1.82	3.82	14.91	31.45	154.12
Finland	10.43	20.69	45.49	1.74	3.61	13.27	27.52	130.87
France	10.11	20.38	45.28	1.73	3.57	13.34	27.97	132.30
German Democratic Republic	10.12	20.33	44.87	1.73	3.56	13.17	27.42	129.92
Federal Republic of Germany	10.16	20.37	44.50	1.73	3.53	13.21	27.61	132.23
Great Britain and Northern Ireland	10.11	20.21	44.93	1.70	3.51	13.01	27.51	129.13
Greece	10.22	20.71	46.56	1.78	3.64	14.59	28.45	134.60
Guatemala	10.98	21.82	48.40	1.89	3.80	14.16	30.11	139.33
Hungary	10.26	20.62	46.02	1.77	3.62	13.49	28.44	132.58
India	10.60	21.42	45.73	1.76	3.73	13.77	28.81	131.98

(d) Conduct a test of $H_0: \Sigma = \mathbf{LL}' + \Psi$ versus $H_1: \Sigma \neq \mathbf{LL}' + \Psi$ for both $m = 2$ and $m = 3$ at the $\alpha = .01$ level. With these results and those in Parts b and c, which choice of m appears to be the best?

(e) Suppose a new salesperson, selected at random, obtains the test scores $\mathbf{x}' = [x_1, x_2, \ldots, x_7] = [110, 98, 105, 15, 18, 12, 35]$. Calculate the salesperson's factor score using the weighted least squares method and the regression method.
Note: The components of $\mathbf{x}$ must be standardized using the sample means and variances calculated from the original data.

9.19. Using the air-pollution variables X_1, X_2, X_5, and X_6 given in Table 1.3, generate the sample *covariance* matrix.

(a) Obtain the principal component solution to a factor model with $m = 1$ and $m = 2$.

(b) Find the maximum likelihood estimates of $\mathbf{L}$ and Ψ for $m = 1$ and $m = 2$.

(c) Compare the factorization obtained by the principal component and maximum likelihood methods.

9.20. Perform a varimax rotation of both $m = 2$ solutions in Exercise 9.19. Interpret the results. Are the principal component and maximum likelihood solutions consistent with each other?

9.21. Refer to Exercise 9.19.

(a) Calculate the factor scores from the $m = 2$ maximum likelihood estimates by (i) weighted least squares in (9-50) and by (ii) the regression approach of (9-58).

(b) Find the factor scores from the principal component solution using (9-51).

(c) Compare the three sets of factor scores.

9.22. Repeat Exercise 9.19 starting from the sample *correlation* matrix. Interpret the factors for the $m = 1$ and $m = 2$ solutions. Does it make a difference if $\mathbf{R}$, rather than $\mathbf{S}$, is factored? Explain.

9.23. Perform a factor analysis of the census-tract data in Table 8.2. Start with $\mathbf{R}$ and obtain both the maximum likelihood and principal component solutions. Comment on your choice of m. Your analysis should include factor rotation and the computation of factor scores.

9.24. Perform a factor analysis of the "stiffness" measurements given in Table 4.3 and discussed in Example 4.13. Compute factor scores and check for outliers in the data. Use the sample covariance matrix $\mathbf{S}$.

9.25. Consider the mice-weight data in Example 8.6. Start with the sample *covariance* matrix (see Exercise 8.15 for $\sqrt{s_{ii}}$).

(a) Obtain the principal component solution to the factor model with $m = 1$ and $m = 2$.

(b) Find the maximum likelihood estimates of the loadings and specific variances for $m = 1$ and $m = 2$.

(c) Perform a varimax rotation of the solutions in Parts a and b.

9.26. Repeat Exercise 9.25 by factoring $\mathbf{R}$ instead of the sample covariance matrix $\mathbf{S}$. Also, for the mouse with standardized weights $[.8, -.2, -.6, 1.5]$, obtain the factor scores using the maximum likelihood estimates of the loadings and Equation (9-58).

9.27. Perform a factor analysis of the national track records for women data given in Table 1.7. Use the sample covariance matrix $\mathbf{S}$ and interpret the factors. Compute factor scores and check for outliers in the data. Repeat the analysis with the sample correlation matrix $\mathbf{R}$. Does it make a difference if $\mathbf{R}$, rather than $\mathbf{S}$, is factored? Explain.

9.28. Perform a factor analysis of the national track records for men data given in Table 8.3. Repeat the steps given in Exercise 9.27. Is the appropriate factor model for the men's data different from the one for the women's data? If not, are the interpretations of the factors roughly the same? If the models are different, explain the differences.