(1, 1pt) 7. It equals the trace of the correlation matrix.

(2, 1pt) $\sqrt{0.04 \times 7} = 0.529$

(3, 1pt) Two, because the first two PCs explain 0.83+0.09=92% of the total variation.

(4, 2pts) The first PC is essentially an average of the normalized track records and might measure the athletic excellence of a given nation (larger score, better athletic performance). The second PC might meaure the relative strength of a nation at the various running distances, basically the contrast between performance of long-distance and short-distance races (large positive score, long better than short; large negative, short better than long; score close to zero, equivalent)

(5, 2pts) Marathon, because its variance (1825.77^2) dominates the total variance (= Var(m100) + $\cdots$ + Var(Marathon) = $0.45^2 + \cdots + 1825.77^2$), and the first PC is the linear combination with maximum variance.

(6, 1pt) The last (smallest) eigenvalue usually indicate the strongest linear dependency among variables. In this case, because 0.15 is quite small, we can say

$$(-0.109) \cdot \frac{\text{m200} - 11.6}{0.45} + \cdots + (-0.598) \cdot \frac{\text{m3000} - 566.9}{49.46} \approx \text{a constant}$$

(7, 2pts) Let S be the sample covariance matrix, R the correlation matrix, and D a diagonal matrix with diagonal entry being $diag(S)$, then $S = D^{1/2}RD^{1/2}$ and

$$|S| = |D^{1/2}||R||D^{1/2}| = |D||R| = \left(\prod_{i=1}^{7} \text{sample variance of variable}_i\right)\left(\prod_{i=1}^{7} \text{sample variance of } \ i\text{th PC}\right)$$

(8, 1pt) zero. Let $\hat{PC}_1 = \hat{e}_1 X$, $\hat{PC}_2 = \hat{e}_2 X$, then

$$cov(\hat{PC}_1, \hat{PC}_2) = \hat{e}_1^T R \hat{e}_2 = 0.$$

Note that if you say it is zero because $cov(PC_1, PC_2) = 0$, your answer is not accurate enough. For example, consider the case of factor analysis.

(9, 2pts) Because the analysis is applied on correlation matrix, we need to standerdize the raw data before calculating the score. The score is

$$\sum_{j=1}^{k} a_{ij}\left(\frac{x_{ij} - \bar{x}_j}{\sqrt{s_{jj}}}\right) = (-0.368)\left(\frac{11.22 - 11.6}{0.45}\right) + \cdots + (-0.367)\left(\frac{10672 - 10395.2}{1825.77}\right)$$

(10, 2pts) I wiould use 1st PC scores because of two reasons. First, the 1st PC has largest variance, which means we can best distinguish the 55 nations using the score. Second, it represnts average athletic performance (lager the better).

(11, 1pt) gdr, ussr, and usa

(12, 2pts) From the 1st PC score, Taiwan is moderate on the athletic excellence. From the 2nd PC score, the contrast between short and long distance is quite large (short-distance better than long-distance), which indicates some kind of unbalance.

(13, 1pt) wsamoa. It is far away from (0, 0) in the plot and the statistical distance calculated from the raw data will be similar to the statistical distance based on the first 2 PC scores because the first 2 PCs explain 92% of total variation.

(14, 1pt) three. Let m be the number of common factors. Then, $LL' + \Psi$ has $7 \times (3 + 1) = 28$ parameters when $m = 3$ and the unrestricted covariance matrix also has $7 \times (7 + 1)/2 = 28$ parameters.

(15, 1pt) $28 - 28 + (3 \times (3 - 1)/2) = 3$

(16, 1pt) normality

(17, 1pt) the communality of m100 is $0.404^2 + 0.833^2 + 0.252^2 = 0.92$

(18, 2pts) 1st common factor represente the long-distance performance and the 2nd mainly concerns with short-distance performance

(19, 2pts) $(0.404, 0.833, 0.252)(0.777, 0.398, 0.248)^T + (-0.022) = 0.686$

(20, 1pt) We can use the R output of PCA to get the answer: $(-0.368) \times 2.41 = -0.887$.

(21, 2pts) We can identify two groups in the plot: {m1500, m3000, marathon} and {m100, m200}. The rotation should make the loading best distinguish the 2 groups. For orthogonal rotation, the 1st axis should be close to group1 and the 2nd axis close to group2 and the two axes are orthogonal. For oblique rotation, the two axes should pass the center of the two groups.