Answers to 2009 STA 437/1005 mid-term test

1a) $\bar{x} \;=\; [100\ 200]'$.

1b)

$x_{j1}-\bar{x}_1$	$(x_{j1}-\bar{x}_1)^2$	$x_{j2}-\bar{x}_2$	$](x_{j2}-\bar{x}_2)^2$	$(x_{j1}-\bar{x}_{j1})(x_{j2}-\bar{x}_2)$
-5	25	10	100	-50
0	0	0	0	0
5	25	-10	100	-50
-5	25	-5	25	25
5	25	5	25	25

$$s_{11} \;=\; (1/4)\textstyle\sum_j (x_{j1}-\bar{x}_1)^2 \;=\; 100/4$$
$$s_{22} \;=\; (1/4)\textstyle\sum_j (x_{j2}-\bar{x}_j)^2 \;=\; 250/4$$
$$s_{12} \;=\; s_{21} \;=\; (1/4)\textstyle\sum_j (x_{j1}-\bar{x}_{j1})(x_{j2}-\bar{x}_2) \;=\; -50/4$$
$$S \;=\; \left[\begin{array}{rr} 100/4 & -50/4 \\ -50/4 & 250/4 \end{array}\right] \;=\; (50/4)\left[\begin{array}{rr} 2 & -1 \\ -1 & 5 \end{array}\right]$$

1c) $$S^{-1} \;=\; (4/50)\left[\begin{array}{rr} 5/9 & 1/9 \\ 1/9 & 2/9 \end{array}\right] \;=\; (4/450)\left[\begin{array}{rr} 5 & 1 \\ 1 & 2 \end{array}\right]$$

$$d_1^2 \;=\; (x_1-\bar{x})'S^{-1}(x_1-\bar{x}) \;=\; [-5\ 10](4/450)\left[\begin{array}{rr} 5 & 1 \\ 1 & 2 \end{array}\right]\left[\begin{array}{r} -5 \\ 10 \end{array}\right] \;=\; 2$$

2a) By the spectral decomposition theorem, we can write any symmetric real matrix, A, as $A \;=\; \lambda_1 e_1 e_1' + \cdots + \lambda_k e_k e_k'$. Since the e_i are orthogonal (hence linearly independent), we can write any non-zero vector, c, as $c \;=\; c_1 e_1 + \cdots + c_k e_k$, for some scalars c_i, not all of which are zero. We can now write $c'Ac$ as follows:

$$c'Ac \;=\; (c_1e_1 + \cdots + c_ke_k)'\,(\lambda_1 e_1 e_1' + \cdots \lambda_k e_k e_k')\,(c_1e_1 + \cdots + c_ke_k)$$

Multiplying this out, and noting that $e_i'e_i = 1$, and also that $e_i'e_j = 0$ when $i \neq j$, we see that

$$c'Ac \;=\; c_1^2\lambda_1 + \cdots + c_k^2\lambda_k$$

If all the λ_k are positive, this must be greater than zero, since at least one of the c_i is non-zero. The matrix A is therefore positive definite.

2b) First, note that the converse of part (a) is true: if a matrix A is positive definite, all its eigenvalues are positive. To see this, suppose e is an eigenvalue of A (which we can take to be of length one), with eigenvalue λ. Then $e'Ae > 0$ by the definition of positive definiteness. But $e'Ae = e'\lambda e = \lambda e'e = \lambda$, so $\lambda > 0$.

Next, write A and B as follows, using the spectral decomposition theorem:

$$A \;=\; \alpha_1 e_1 e_1' \;+\; \cdots \;+\; \alpha_k e_k e_k'$$
$$B \;=\; \beta_1 e_1 e_1' \;+\; \cdots \;+\; \beta_k e_k e_k'$$

where $e_1, \ldots, e_k$ are the eigenvectors of both A and B, $\alpha_1, \ldots, \alpha_k$ are the eigenvalues of A, and $\beta_1, \ldots, \beta_k$ are the eigenvalues of B.

When we now multiply A and B, and note that $e_i'e_i = 1$, and also that $e_i'e_j = 0$ when $i \neq j$, we get

$$AB \;=\; \alpha_1\beta_1 e_1 e_1' + \cdots + \alpha_k\beta_k e_k e_k'$$

Since $(e_i e_i')' = e_i'' e_i' = e_i e_i'$, each term above is symmetric, so AB is also symmetric. We can easily see that the eigenvectors of AB are the same e_i, and that the eigenvalues are the $\alpha_i \beta_i$ — for instance

$$ABe_1 = \alpha_1\beta_1 e_1 e_1' e_1 + \alpha_2\beta_2 e_2 e_2' e_1 + \cdots = \alpha_1\beta_1 e_1$$

If A and B are positive definite, then all the α_i and β_i are positive, and therefore all the eigenvalues of AB are positive. By part (a), AB is therefore positive definite.

3a) A

3b) B

3c) B

3d) C

The T^2 statistic for testing $H_0 : \ \mu = 0$ versus $H_1 : \ \mu \neq 0$ is

$$T^2 = n\bar{x}' S_X^{-1} \bar{x}$$

Let A be any non-singular $p \times p$ matrix. Transforming to $y = Ax$, we will find that the sample mean is $\bar{y} = A\bar{x}$ and the sample covariance matrix is $S_Y = AS_X A'$. The new T^2 statistic will then be

$$\begin{aligned} T^2 &= n\bar{y}' S_Y^{-1} \bar{y} \\ &= n(A\bar{x})'(AS_X A')^{-1}(A\bar{x}) \\ &= n\bar{x}' A' A'^{-1} S_X^{-1} A^{-1} A\bar{x} \\ &= n\bar{x}' S_X^{-1} \bar{x} \end{aligned}$$

which is the same as the old T^2 statistic.

4a) D, no outliers, not normal (heavy-tailed).

4b) E, no outliers, not normal (light-tailed).

4c) A, no outliers, no reason to think it's not normal (note the small sample size).

4d) B, one outlier (rightmost point, about 7), no reason to think it's not normal (after ignoring the outlier).

5a)

$$\Sigma = \begin{bmatrix} 5 & 1 & 1 & 1 \\ 1 & 6 & 2 & 2 \\ 1 & 2 & 7 & 3 \\ 1 & 2 & 3 & 8 \end{bmatrix}$$

5b)

$$\text{Var}(\overline{Y}) = [1/4 \ 1/4 \ 1/4 \ 1/4] \ \Sigma \begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = 23/8$$

5c) Normal, with mean $0 \ + \ 1(8)^{-1}(y_4 - 0) \ = \ y_4/8$ and variance $5 \ - \ 1(8)^{-1}1 \ = \ 39/8$.