1. Define the trace of a square matrix.

2. Assuming the product AB is square, show that tr(AB)=tr(BA).

3. Given an (nxk) matrix X, and an (nxn) diagonal matrix W:
   a. Show X'X is p.s.d..
   b. Show X'X is p.d. if X has rank k.
   c. Show X'WX is p.d. if X has rank k and W_{ii}>0 for all i.
   d. Is the matrix X'X symmetric?
   e. Is the matrix X(X'X)^{-1}X' symmetric?
   f. Let M=I_n-X(X'X)^{-1}X'. Show that MM=M.

4. Show that the rank of an idempotent matrix equals its trace.

5. Show that if A is an (nxn) p.d. matrix, then an (nxn) non-singular matrix B may be found, such that A=BB'.

6. State the rule for partitioning a pair of matrices to conform for addition.

7. State the rule for partitioning a pair of matrices to conform for multiplication.

8. Let X denote an (nxk) matrix, W an (nxn) diagonal matrix, and Y an (nxm) matrix. Show that
   \[ X'WY = \sum_{i=1}^{n} W_{ii} X_i' Y_i \]
   where X_i and Y_i denote the rows of X and Y.


10. Use the formula for the inverse of a partitioned matrix to show that the inverse of a block diagonal matrix is found by inverting the diagonal blocks.

11. Let X and Y denote (nx1) vectors, W and Z scalars, and A an nxn matrix. Find the following partial derivatives:
   a. \( \partial (X'Y) / \partial Y \) and \( \partial (X'Y) / \partial X \)
   b. \( \partial (X'AX) / \partial X \) and \( \partial (X'AX) / \partial A \)
   c. \( \partial Z / \partial X' \) where Z=f(W) and W=g(X) are differentiable functions
   d. \( \partial Z / \partial X' \) where Z=f(Y) and Y=g(X) are differentiable functions
1. The trace of an \((nxn)\) matrix \(A\) is defined as the sum of the diagonal elements. That is,
\[
\text{tr}(A) = \sum_{i=1}^{n} a_{ii}.
\]

2. Since \(AB\) is square, we may assume without loss in generality, that \(A\) is \((nxm)\) and \(B\) is \((mxn)\). Let \(C=AB\) and \(D=BA\), then
\[
c_{ij} = \sum_{h=1}^{m} a_{ih} b_{hj} \quad \text{for } i,j = 1, \ldots, n
\]
and
\[
d_{hk} = \sum_{i=1}^{n} b_{hi} a_{ik} \quad \text{for } h,k = 1, \ldots, m.
\]
Thus,
\[
\text{tr}(AB) = \text{tr}(C) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n} [\sum_{h=1}^{m} a_{ih} b_{hi}]
\]
and
\[
\text{tr}(BA) = \text{tr}(D) = \sum_{h=1}^{m} d_{hh} = \sum_{i=1}^{n} [\sum_{h=1}^{m} b_{hi} a_{ih}]
\]
Thus, \(\text{tr}(AB) = \text{tr}(BA)\).

3a. The matrix \(X'X\) is p.s.d. if the quadratic form \(Y'(X'X)Y \geq 0\) for any \(Y \neq 0\). In this case, \(Y\) is an arbitrary \(k\)-vector. But \(Y'(X'X)Y\) may be written as \(Z'Z\), where \(Z=XY\) is an \(n\)-vector. Note that when \(X\) does not have full column rank, \(k\), we may have \(Z=0\) even if \(Y \neq 0\). Thus, \(Y'(X'X)Y=Z'Z=\sum_{i=1}^{n} z_{i}^2 \geq 0\) for any \(Y \neq 0\), and \(X'X\) is p.s.d..

3b. The matrix \(X'X\) is p.d. if the quadratic form \(Y'(X'X)Y > 0\) for any \(Y \neq 0\). As above, we may write \(Y'(X'X)Y=Z'Z=\sum_{i=1}^{n} z_{i}^2 > 0\) for any \(Y \neq 0\), since \(Y \neq 0\) implies \(Z \neq 0\) if \(X\) has rank \(k\). (By definition, \(X\) has rank \(k\) if \(XY\neq 0\) for any \(Y \neq 0\).) Thus, \(X'X\) is p.d. when \(X\) has rank \(k\).

3c. The matrix \(X'WX\) is p.d. if the quadratic form \(Y'(X'WX)Y > 0\) for any \(Y \neq 0\). Since \(W\) is a diagonal matrix, we may write \(Y'(X'WX)Y=Z'WZ=\sum_{i=1}^{n} w_{ii} z_{i}^2 > 0\) for any \(Y \neq 0\), since \(Y \neq 0\) implies \(Z \neq 0\) if \(X\) has rank \(k\), and since \(w_{ii} > 0\) for all \(i\). Thus, \(X'WX\) is p.d. when \(X\) has rank \(k\) and \(w_{ii} > 0\) for all \(i\).

3d. The matrix \(X'X\) is symmetric if \((X'X)'=(X'X)\). But \((X'X)'=X'(X')'=X'X\) since \((AB)'=B'A'\) and \((A')'=A\). Thus, \(X'X\) is symmetric.

3e. \([X(X'X)^{-1}X]'=[(X'X)^{-1}]X'=X[(X'X)^{-1}]X'\) since \((AB)'=B'A'\), \((A')'=A\), and \((A^{-1})'=(A')^{-1}\). Thus, \((X'X)^{-1}X'\) is symmetric.

3f. \(MM=[I-X(X'X)^{-1}X']^2[I-X(X'X)^{-1}X'] =I-X(X'X)^{-1}X'-X(X'X)^{-1}X'X(X'X)^{-1}X'\)
\(=I-X(X'X)^{-1}X'-X(X'X)^{-1}X'+X(X'X)^{-1}X'\) since \(X(X'X)^{-1}=I\)
\(=I-X(X'X)^{-1}X'=M\).
4. Let $V$ denote the characteristic vectors of a matrix $M$. Then, we know that:
   a. $V'MV=\text{diag}(r_1, \ldots, r_n)=R$, where $r_i$ denotes the $i^{th}$ characteristic root of $M$. That is, the characteristic vectors diagonalize the matrix $M$.
   b. $V'V=VV'=I$. That is, characteristic vectors are orthonormal.
   c. If $M$ is idempotent matrix, then for all $i$, either $r_i=1$ or $r_i=0$.
   d. The rank of $M$ is equal to the number of non-zero characteristic roots.

Properties (a) and (b) imply that $M=VRV'$. Since $\text{tr}(AB)=\text{tr}(BA)$, $\text{tr}(M)=\text{tr}(VRV')=\text{tr}(RV'V)=\text{tr}(R)$. Given property (c), we have $\text{tr}(M)=\sum_{i=1}^{n} r_i$. Finally, given property (d), $\text{tr}(M)=\rho(M)$.

5. Properties (a) and (b) above imply that $A=VRV'$, where $V$ and $R$ are based upon the characteristic roots and vectors of $A$. Since $A$ is p.d., its characteristic roots must all be strictly positive. Hence, we may factor $R$ into the product $PP'$, where $P=\text{diag}\sqrt{r_1, \ldots, r_n}$. The matrix $A$ may now be written as $A=VRV'=VPP'V'=BB'$ where $B=VP$. Since $B$ is the product of non-singular matrices, it is also non-singular.

6. In order for a pair of partitioned matrices to conform for addition, the entire matrices must be of the same dimension and corresponding submatrices must also be of the same dimension.

7. In order for a pair of partitioned matrices to conform for multiplication, the column dimension of the premultiplier must equal the row dimension of the postmultiplier, and the same partition must be applied to the columns of the premultiplier and the rows of the postmultiplier.

8. Since $W$ is diagonal, the product $WY$ is composed of the $m$-dimensional row vectors $[W_{ii}Y_i]$, $i=1,\ldots,n$, where $Y_i$ denotes the $i^{th}$ row of $Y$. The product $X'WY$ is then

$$X'WY = \begin{bmatrix} W_{11}Y_1 \\ \vdots \\ W_{nn}Y_n \end{bmatrix} = \begin{bmatrix} X'_1 \\ \vdots \\ X'_n \end{bmatrix}$$

which gives $X'WY=\sum_{i=1}^{n} W_{ii}X'_iY_i$.

9. A block diagonal matrix is a square matrix which may be written as a partitioned matrix with square diagonal blocks and zero off diagonal blocks.
10. Using the formula for partitioned inversion, the diagonal blocks of $A^{-1}$ are

$$(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} \quad \text{and} \quad A_{22}^{-1}+A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1}$$

which reduce to $A_{11}^{-1}$ and $A_{22}^{-1}$, respectively, since $A_{12}=0$ and $A_{21}=0$ if $A$ is block diagonal. The off-diagonal blocks of $A^{-1}$ are

$$-(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \quad \text{and} \quad -A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}$$

which reduce to zero since $A_{12}=0$ and $A_{21}=0$ if $A$ is block diagonal.

11a. The vector product $X'Y=\Sigma_{i=1}^{n}X_iY_i$. Hence, the (nx1) vector of partial derivatives $\partial(X'Y)/\partial Y$ is composed of the elements $\partial(X'Y)/\partial Y_j=X_j$, for $j=1,...,n$. This gives $\partial(X'Y)/\partial Y=X$. Similarly, the (nx1) vector of partial derivatives $\partial(X'Y)/\partial X$ is composed of the elements $\partial(X'Y)/\partial X_j=Y_j$, for $j=1,...,n$. This gives $\partial(X'Y)/\partial X=Y$.

11b. The quadratic form $X'AX=\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}a_{ij}X_iX_j$. Hence, the (nx1) vector of partial derivatives $\partial(X'AX)/\partial X$ is composed of the elements $\partial(X'AX)/\partial X_k=\Sigma_{i=1}^{n}a_{ik}X_i+\Sigma_{j=1}^{n}a_{kj}X_j=A_k'X+A_kX$ for $k=1,...,n$. In this case, $A_k$ denotes the $k^{th}$ row of $A$. This gives $\partial(X'AX)/\partial X=A'X+AX$, and if $A$ is symmetric, $\partial(X'AX)/\partial X=2AX$. Similarly, the (nxn) vector of partial derivatives $\partial(X'AX)/\partial A$ is composed of the elements $\partial(X'AX)/\partial a_{kh}=X_kX_h$ for $k,h=1,...,n$. This gives $\partial(X'AX)/\partial A=XX'$.

11c. This is a composite function, $Z=f(g(X))$, where $Z$ and $W=g(X)$ are scalars and $X$ is an n-vector. By the chain rule, $\partial Z/\partial X_j = \partial Z/\partial W \partial W/\partial X_j$. Thus, $\partial Z/\partial X' = \partial Z/\partial W \partial W/\partial X'$.

11d. This is also a composite function, $Z=f(g(X))$, where $Y=g(X)$ and $X$ are n-vectors and $Z$ is a scalar. By the chain rule, $\partial Z/\partial X_j = \Sigma_{i=1}^{n} \partial Z/\partial Y_i \partial Y_i/\partial X_j$. Thus, $\partial Z/\partial X' = \partial Z/\partial Y' \partial Y/\partial X'$.

References

