Appendix A: Fun Matrix Facts

For convenience, we summarize several useful matrix facts here.

**SQUARE MATRIX RESULTS:** Let \( A \) and \( B \) be square matrices of the same dimension. Inverses below are assumed to exist.

- \((AB)^{-1} = B^{-1}A^{-1}\), \((A^{-1})^T = (A^T)^{-1}\).
- We denote the **determinant** of \( A \) by \( |A| \).
- \(|A| = |A^T|, |A| = 1/|A^{-1}|\)
- \(|AB| = |A||B|\)
- \((A + B)^{-1} = A^{-1} - A^{-1}(A^{-1} + B^{-1})^{-1}A^{-1}\)
- \((A + CBD)^{-1} = A^{-1} - A^{-1}C(B^{-1} + DA^{-1}C)^{-1}DA^{-1}\). Here, \( A \) and \( B \) need not be of the same dimension, and \( C \) and \( D \) are conformable matrices.
- The following are equivalent: (i) \( A \) is **nonsingular**, (ii) \(|A| \neq 0\), (iii) \( A^{-1} \) exists.
- We denote the **trace** of a square matrix \( A \) by \( \text{tr}(A) \).
- \( \text{tr}(A) = \text{tr}(A^T), \text{tr}(bA) = b\text{tr}(A)\)
- \( \text{tr}(A + B) = \text{tr}(A) + \text{tr}(B), \text{tr}(AB) = \text{tr}(BA)\)
- If \( A \) is \((n \times n)\) and \( x \) is \((n \times 1)\), then the **quadratic form**

\[ x^TAx \]

and \( A \) are **nonnegative definite** if \( x^TAx \geq 0 \). The quadratic form and \( A \) are **positive definite** if \( x^TAx > 0 \). If \( A \) is positive definite, then it is symmetric and nonsingular (so its inverse exists).

- \( x^TAx = \text{tr}(Axx^T)\)
- If \( x \) is a **random vector** with mean \( \mu \) and covariance matrix \( V \), then

\[ E(x^TAx) = \text{tr}\{E(xx^TA)\} = \text{tr}(VA) + \mu^TA\mu = \text{tr}(AV) + \mu^TA\mu. \]
**vec and vech NOTATION:**

- For a \((n \times r)\) matrix \(A\), vec\((A)\) is defined as the \((nr \times 1)\) vector consisting of the \(r\) columns of \(A\) stacked in the order 1, \ldots, \(r\).

- If furthermore \(A\) is \((n \times n)\) and symmetric, then vec\((A)\) contains redundant entries. The vech\((\cdot)\) operator yields the column vector containing all the distinct entries of \(A\) by stacking the lower diagonal elements; e.g., for \(n = 3\),

\[
A = \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
a_{12} & a_{22} & a_{23} \\
a_{13} & a_{23} & a_{33}
\end{pmatrix}
\quad \text{and} \quad
\text{vech}(A) = \begin{pmatrix}
a_{11} \\
a_{12} \\
a_{13} \\
a_{22} \\
a_{23} \\
a_{33}
\end{pmatrix}.
\]

- For matrices \(A\) \((a \times a)\), \(B\), \(C\), \(D\),

  (i) \(\text{tr}(AB) = \{\text{vec}(A)\}^T\{\text{vec}(B^T)\} = \{\text{vec}(A^T)\}^T\{\text{vec}(B)\}\).

  (ii) \(\text{tr}(ABD^TC^T) = \{\text{vec}(A)\}^T(B \otimes C)\text{vec}(D)\), where \(\otimes\) represents Kronecker product.

  (iii) For \(A\) symmetric, there is a relationship between vec\((A)\) and vech\((A)\). In particular, there exists a unique matrix \(\Phi\) of dimension \(\{a^2 \times a(a + 1)/2\}\) such that

\[
\text{vec}(A) = \Phi\text{vech}(A).
\]

Clearly, \(\Phi\) is unique and of full column rank, as there is only one way to write the distinct elements of \(A\) in a full, redundant vector.
**INVERSE OF PARTITIONED MATRIX:** Consider a generic \((k \times k)\) matrix
\[
C = \begin{pmatrix}
    C_{11} & C_{12} \\
    C_{21} & C_{22}
\end{pmatrix},
\]
where the \(C_{ij}\) are submatrices such that \(C_{11}\) is \((k_1 \times k_1)\) and \(C_{22}\) is \((k_2 \times k_2)\) such that \(k = k_1 + k_2\), and \(C_{11}^{-1}\) and \(C_{22}^{-1}\) exist, as do all other inverses below. Then
\[
C^{-1} = \begin{pmatrix}
    D_{11} & D_{12} \\
    D_{21} & D_{22}
\end{pmatrix},
\]
where
\[
D_{11} = (C_{11} - C_{12}C_{22}^{-1}C_{21})^{-1}
\]
\[
D_{22} = (C_{22} - C_{21}C_{11}^{-1}C_{12})^{-1} = C_{22}^{-1} + C_{22}^{-1}C_{21}D_{11}C_{12}C_{22}^{-1}
\]
\[
D_{12} = -C_{11}^{-1}C_{12}D_{22} = -D_{11}C_{12}C_{22}^{-1}
\]
\[
D_{21} = -C_{22}^{-1}C_{21}D_{11}.
\]

**MATRIX DIFFERENTIATION:** Let \(x\) be a \((n \times 1)\) vector depending on a \((p \times 1)\) vector \(\beta\), and let \(A\) be a \((n \times n)\) square matrix.

- For quadratic form \(Q = x^T Ax\), \(\partial Q / \partial x = 2Ax\). Note that this is a \((n \times 1)\) vector.
- The chain rule then gives \(\partial Q / \partial \beta = (\partial x / \partial \beta)(\partial Q / \partial x)\). Note that \(\partial x / \partial \beta\) is a \((p \times n)\) matrix.

Let \(V(\xi)\) be a \((n \times n)\) nonsingular matrix depending on a \((q \times 1)\) (parameter) vector \(\xi\).

- If \(\xi_k\) is the \(k\)th element of \(\xi\), then \(\partial / \partial \xi_k V(\xi)\) is the \((n \times n)\) matrix whose \((\ell, \rho)\) element is the partial derivative of the \((\ell, \rho)\) element of \(V(\xi)\) with respect to \(\xi_k\).
- \(\partial / \partial \xi_k \{\log |V(\xi)|\} = \text{tr} \left[ V^{-1}(\xi) \{\partial / \partial \xi_k V(\xi)\} \right].\)
- \(\partial / \partial \xi_k V^{-1}(\xi) = -V^{-1}(\xi) \{\partial / \partial \xi_k V(\xi)\} V^{-1}(\xi).\)
- For quadratic form \(Q = x^T V(\xi)x\), \(\partial Q / \partial \xi_k = x^T \{\partial / \partial \xi_k V(\xi)\} x\). Thus, from the previous result,
\[
\partial / \partial \xi_k \{x^T V^{-1}(\xi)x\} = -x^T V^{-1}(\xi) \{\partial / \partial \xi_k V(\xi)\} V^{-1}(\xi)x.
\]