Appendix A: Fun Matrix Facts

For convenience, we summarize several useful matrix facts here.

SQUARE MATRIX RESULTS: Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be square matrices of the same dimension. Inverses below are assumed to exist.

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

  $$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$$

  and $\boldsymbol{A}$ are ***nonnegative definite*** if $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} \geq 0$. The quadratic form and $\boldsymbol{A}$ are ***positive definite*** if $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} > 0$. If $\boldsymbol{A}$ is positive definite, then it is symmetric and nonsingular (so its inverse exists).
- $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \mathrm{tr}(\boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^T)$
- If $\boldsymbol{x}$ is a ***random vector*** with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{V}$, then

  $$E(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}) = \mathrm{tr}\{E(\boldsymbol{x}\boldsymbol{x}^T)\boldsymbol{A}\} = \mathrm{tr}(\boldsymbol{VA}) + \boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu} = \mathrm{tr}(\boldsymbol{AV}) + \boldsymbol{\mu}^T\boldsymbol{A}\boldsymbol{\mu}.$$

vec and vech NOTATION:

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

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

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

- For matrices $\boldsymbol{A}$ $(a \times a)$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$,

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

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

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

  $$\text{vec}(\boldsymbol{A}) = \boldsymbol{\Phi}\text{vech}(\boldsymbol{A}).$$

  Clearly, $\boldsymbol{\Phi}$ is unique and of full column rank, as there is only one way to write the distinct elements of $\boldsymbol{A}$ in a full, redundant vector.

INVERSE OF PARTITIONED MATRIX: Consider a generic $(k \times k)$ matrix

$$\boldsymbol{C} = \begin{pmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{pmatrix},$$

where the $\boldsymbol{C}_{ij}$ are submatrices such that $\boldsymbol{C}_{11}$ is $(k_1 \times k_1)$ and $\boldsymbol{C}_{22}$ is $(k_2 \times k_2)$ such that $k = k_1 + k_2$, and $\boldsymbol{C}_{11}^{-1}$ and $\boldsymbol{C}_{22}^{-1}$ exist, as do all other inverses below. Then

$$\boldsymbol{C}^{-1} = \begin{pmatrix} \boldsymbol{D}_{11} & \boldsymbol{D}_{12} \\ \boldsymbol{D}_{21} & \boldsymbol{D}_{22} \end{pmatrix},$$

where

$$\begin{aligned} \boldsymbol{D}_{11} &= (\boldsymbol{C}_{11} - \boldsymbol{C}_{12}\boldsymbol{C}_{22}^{-1}\boldsymbol{C}_{21})^{-1} \\ \boldsymbol{D}_{22} &= (\boldsymbol{C}_{22} - \boldsymbol{C}_{21}\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12})^{-1} = \boldsymbol{C}_{22}^{-1} + \boldsymbol{C}_{22}^{-1}\boldsymbol{C}_{21}\boldsymbol{D}_{11}\boldsymbol{C}_{12}\boldsymbol{C}_{22}^{-1} \\ \boldsymbol{D}_{12} &= -\boldsymbol{C}_{11}^{-1}\boldsymbol{C}_{12}\boldsymbol{D}_{22} = -\boldsymbol{D}_{11}\boldsymbol{C}_{12}\boldsymbol{C}_{22}^{-1} \\ \boldsymbol{D}_{21} &= -\boldsymbol{C}_{22}^{-1}\boldsymbol{C}_{21}\boldsymbol{D}_{11}. \end{aligned}$$

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

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

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

- If ξ_k is the kth element of $\boldsymbol{\xi}$, then $\partial/\partial\xi_k \, \boldsymbol{V}(\boldsymbol{\xi})$ is the $(n \times n)$ matrix whose (ℓ, p) element is the partial derivative of the (ℓ, p) element of $\boldsymbol{V}(\boldsymbol{\xi})$ with respect to ξ_k.
- $\partial/\partial\xi_k\{\log|\boldsymbol{V}(\boldsymbol{\xi})|\} = \mathrm{tr}\left[\boldsymbol{V}^{-1}(\boldsymbol{\xi})\{\partial/\partial\xi_k \boldsymbol{V}(\boldsymbol{\xi})\}\right]$.
- $\partial/\partial\xi_k \boldsymbol{V}^{-1}(\boldsymbol{\xi}) = -\boldsymbol{V}^{-1}(\boldsymbol{\xi})\,\{\partial/\partial\xi_k \, \boldsymbol{V}(\boldsymbol{\xi})\}\,\boldsymbol{V}^{-1}(\boldsymbol{\xi})$.
- For quadratic form $Q = \boldsymbol{x}^T\boldsymbol{V}(\boldsymbol{\xi})\boldsymbol{x}$, $\partial Q/\partial\xi_k = \boldsymbol{x}^T\{\partial/\partial\xi_k \, \boldsymbol{V}(\boldsymbol{\xi})\}\boldsymbol{x}$. Thus, from the previous result,

$$\partial/\partial\xi_k\,\{\boldsymbol{x}^T\boldsymbol{V}^{-1}(\boldsymbol{\xi})\boldsymbol{x}\} = -\boldsymbol{x}^T\boldsymbol{V}^{-1}(\boldsymbol{\xi})\,\{\partial/\partial\xi_k \, \boldsymbol{V}(\boldsymbol{\xi})\}\,\boldsymbol{V}^{-1}(\boldsymbol{\xi})\boldsymbol{x}.$$