Methods 2 - 6

Chris Mathys

BSc Programme in Cognitive Science

Spring 2023

Linear Algebra: Vectors and Matrices – Gill, chapter 4

- **Matrices**
  - Trace and determinant
  - Rank
  - Norms
  - Matrix inversion

Properties of (Conformable) Matrix Trace Operations

$\rightarrowtail$	Identity Matrix	$\text{tr}(\mathbf{I}_n) = n$
$\rightarrowtail$	Zero Matrix	$\text{tr}(\mathbf{0}) = 0$
$\rightarrowtail$	Square $\mathbf{J}$ Matrix	$\text{tr}(\mathbf{J}_n) = n$
$\rightarrowtail$	Scalar Multiplication	$\text{tr}(s\mathbf{X}) = s\text{tr}(\mathbf{X})$
$\rightarrowtail$	Matrix Addition	$\text{tr}(\mathbf{X} + \mathbf{Y}) = \text{tr}(\mathbf{X}) + \text{tr}(\mathbf{Y})$
$\rightarrowtail$	Matrix Multiplication	$\text{tr}(\mathbf{XY}) = \text{tr}(\mathbf{YX})$
$\rightarrowtail$	Transposition	$\text{tr}(\mathbf{X}') = \text{tr}(\mathbf{X})$

Determinant

$$\det(\mathbf{X}) = |\mathbf{X}| = \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = x_{11}x_{22} - x_{12}x_{21}.$$

Determinant

$$\det(\mathbf{X}) = |\mathbf{X}| = \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = x_{11}x_{22} - x_{12}x_{21}.$$

$$\det(\mathbf{X}) = \sum_{j=1}^{n} (-1)^{i+j} x_{ij} |\mathbf{X}_{[ij]}|$$

$$\begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix}$$

$$= (+1)x_{11} \begin{vmatrix} x_{22} & x_{23} \\ x_{32} & x_{33} \end{vmatrix} + (-1)x_{12} \begin{vmatrix} x_{11} & x_{13} \\ x_{31} & x_{33} \end{vmatrix} + (+1)x_{13} \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix}.$$

Properties of $(n \times n)$ Matrix Determinants

$\rightarrowtail$	Diagonal Matrix	$\|\mathbf{D}\| = \prod_{i=1}^{n} \mathbf{D}_{ii}$
$\rightarrowtail$	(Therefore) Identity Matrix	$\|\mathbf{I}\| = 1$
$\rightarrowtail$	Triangular Matrix (upper or lower)	$\|\boldsymbol{\theta}\| = \prod_{i=1}^{n} \boldsymbol{\theta}_{ii}$
$\rightarrowtail$	Scalar Times Diagonal	$\|s\mathbf{D}\| = s^n \|\mathbf{D}\|$
$\rightarrowtail$	Transpose Property	$\|\mathbf{X}\| = \|\mathbf{X}'\|$
$\rightarrowtail$	**J** Matrix	$\|\mathbf{J}\| = 0$

Matrix rank

More specifically, when any one column of a matrix can be produced by nonzero scalar multiples of other columns added, then we say that the matrix is not **full rank** (sometimes called **short rank**). In this case at least one column is **linearly dependent**. This simply means that we can produce the relative relationships defined by this column from the other columns and it thus adds nothing to our understanding of the relationships defined by the matrix. One way to look at this is to say that the matrix in question does not "deserve" its number of columns.

Matrix rank

Conversely, the collection of vectors determined by the columns is said to be **linearly independent** columns if the only set of scalars, $s_1, s_2, \ldots, s_j$, that satisfies $s_1\mathbf{x}_{.1} + s_2\mathbf{x}_{.2} + \cdots + s_j\mathbf{x}_{.j} = \mathbf{0}$ is a set of all zero values, $s_1 = s_2 = \ldots = s_j = 0$. This is just another way of looking at the same idea since such a condition means that we ***cannot*** reproduce one column vector from a linear combination of the others.

Matrix rank

Actually this emphasis on columns is somewhat unwarranted because the rank of a matrix is equal to the rank of its transpose. Therefore, everything just said about columns can also be said about rows. To restate, ***the row rank of any matrix is also its column rank***. This is a very important result and is proven in virtually every text on linear algebra. What makes this somewhat confusing is additional terminology. An $(i \times j)$ matrix is **full column rank** if its rank equals the number of columns, and it is **full row rank** if its rank equals its number of rows. Thus, if $i > j$, then the matrix can be full column rank but never full row rank. This does not necessarily mean that it ***has*** to be full column rank just because there are fewer columns than rows.

Matrix rank

It should be clear from the example that a (square) matrix is full rank if and only if it has a nonzero determinant. This is the same thing as saying that a matrix is full rank if it is nonsingular or invertible (see Section 4.6 below). This is a handy way to calculate whether a matrix is full rank because the linear dependency within can be subtle (unlike our example above). In the next section we will explore matrix features of this type.

Properties of Matrix Rank

$\rightarrowtail$	Transpose	$\text{rank}(\mathbf{X}) = \text{rank}(\mathbf{X}')$
$\rightarrowtail$	Scalar Multiplication (nonzero scalars)	$\text{rank}(s\mathbf{X}) = \text{rank}(\mathbf{X})$
$\rightarrowtail$	Matrix Addition	$\text{rank}(\mathbf{X} + \mathbf{Y}) \leq \text{rank}(\mathbf{X}) + \text{rank}(\mathbf{Y})$
$\rightarrowtail$	Consecutive Blocks	$\text{rank}[\mathbf{X}\mathbf{Y}] \leq \text{rank}(\mathbf{X}) + \text{rank}(\mathbf{Y})$ $\text{rank}\left[\begin{smallmatrix}\mathbf{X}\\ \mathbf{Y}\end{smallmatrix}\right] \leq \text{rank}(\mathbf{X}) + \text{rank}(\mathbf{Y})$
$\rightarrowtail$	Diagonal Blocks	$\text{rank}\left[\begin{smallmatrix}\mathbf{X} & 0\\ 0 & \mathbf{Y}\end{smallmatrix}\right] = \text{rank}(\mathbf{X}) + \text{rank}(\mathbf{Y})$
$\rightarrowtail$	Kronecker Product	$\text{rank}(\mathbf{X} \otimes \mathbf{Y}) = \text{rank}(\mathbf{X})\text{rank}(\mathbf{Y})$

Properties of Matrix Norms, Size $(i \times j)$

$\rightarrowtail$	Constant Multiplication	$\|k\mathbf{X}\| = \|k\|\|\mathbf{X}\|$
$\rightarrowtail$	Addition	$\|\mathbf{X}+\mathbf{Y}\| \leq \|\mathbf{X}\| + \|\mathbf{Y}\|$
$\rightarrowtail$	Vector Multiplication	$\|\mathbf{X}\mathbf{v}\|_p \leq \|\mathbf{X}\|_p\|\mathbf{v}\|_p$
$\rightarrowtail$	Norm Relation	$\|\mathbf{X}\|_2 \leq \|\mathbf{X}\|_F \leq \sqrt{j}\|\mathbf{X}\|_2$
$\rightarrowtail$	Unit Vector Relation	$\mathbf{X}'\mathbf{X}\mathbf{v} = (\|\mathbf{X}\|_2)^2\mathbf{v}$
$\rightarrowtail$	P-norm Relation	$\|\mathbf{X}\|_2 \leq \sqrt{\|\mathbf{X}\|_1 \; \|\mathbf{X}\|_\infty}$
$\rightarrowtail$	Schwarz Inequality	$\|\mathbf{X}\cdot\mathbf{Y}\| \leq \|\mathbf{X}\| \; \|\mathbf{Y}\|$, where $\|\mathbf{X}\cdot\mathbf{Y}\| = \mathrm{tr}(\mathbf{X}'\mathbf{Y})$

Matrix inversion

Just like scalars have inverses, some ***square*** matrices have a **matrix inverse**.

The inverse of a matrix $\mathbf{X}$ is denoted $\mathbf{X}^{-1}$ and defined by the property

$$\mathbf{X}\mathbf{X}^{-1} = \mathbf{X}^{-1}\mathbf{X} = \mathbf{I}.$$

Matrix inversion

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$

$$\mathbf{X}^{-1} = \det(\mathbf{X})^{-1} \begin{bmatrix} x_{22} & -x_{12} \\ -x_{21} & x_{11} \end{bmatrix}.$$

Properties of $n \times n$ Nonsingular Matrix Inverse

$\rightarrowtail$	Diagonal Matrix	$\mathbf{D}^{-1}$ has diagonal values $1/d_{ii}$ and zeros elsewhere.
$\rightarrowtail$	(Therefore) Identity Matrix	$\mathbf{I}^{-1} = \mathbf{I}$
$\rightarrowtail$	(Non-zero) Scalar Multiplication	$(s\mathbf{X})^{-1} = \frac{1}{s}\mathbf{X}^{-1}$
$\rightarrowtail$	Iterative Inverse	$(\mathbf{X}^{-1})^{-1} = \mathbf{X}$
$\rightarrowtail$	Exponents	$\mathbf{X}^{-n} = (\mathbf{X}^{n})^{-1}$
$\rightarrowtail$	Multiplicative Property	$(\mathbf{XY})^{-1} = \mathbf{Y}^{-1}\mathbf{X}^{-1}$
$\rightarrowtail$	Transpose Property	$(\mathbf{X}')^{-1} = (\mathbf{X}^{-1})'$
$\rightarrowtail$	Orthogonal Property	If $\mathbf{X}$ is orthogonal, then $\mathbf{X}^{-1} = \mathbf{X}'$
$\rightarrowtail$	Determinant	$\lvert\mathbf{X}^{-1}\rvert = 1/\lvert\mathbf{X}\rvert$