Chapter 0

Matrix Algebra Review

0.1 Matrix Axioms

0.1.1 Matrix Algebra

Matrices are rectangular arrays of numbers. If A is a matrix with dimensions $m \times n$, it is an array with m rows and n columns. We use the element-wise shorthand

$$A \;=\; (a_{ij})$$

to indicate that the elements of A are values $a_{i\,j}$, with $1 \;\leq\; i \;\leq\; m$, $1 \;\leq\; j \;\leq\; n$. (In mathematical linear algebra, subscripts usually start at 1.) **Square matrices** have $m = n$. Matrices form an **algebra**: a set of values on which addition and multiplication are defined:

$$\begin{array}{lcll} C \;=\; A + B & \Leftrightarrow & (c_{ij}) \;=\; (a_{ij}) \;+\; (b_{ij}) & \text{addition} \\ C \;=\; A\,B & \Leftrightarrow & (c_{ij}) \;=\; (\sum_k a_{ik}\, b_{kj}) & \text{multiplication} \end{array}$$

These definitions only make sense when the dimensions of A and B are **conformable**; for addition the dimensions must match, and for multiplication the second dimension of A must match the first dimension of B.

0.1.2 Matrix Axioms

Basic axioms of any algebra:

$$\begin{array}{lcll} A \;+\; B & = & B \;+\; A & \text{Commutativity of addition} \\ (A \;+\; B) \;+\; C & = & A \;+\; (B \;+\; C) & \text{Associativity of addition} \\ (A\,B)\,C & = & A\,(B\,C) & \text{Associativity of multiplication} \\ A\,(B + C) & = & A\,B \;+\; A\,C & \text{Distribution of multiplication over addition} \\ (B + C)\,A & = & B\,A + C\,A & \end{array}$$

For matrices we can demonstrate these axioms using element-wise notation. For example, with matrices the final property (associativity of matrix multiplication) can be derived by expanding the sums in the products:

$$\begin{aligned} (A\,B)\,C \;\; &= \;\; \left(\sum_k a_{ik}\, b_{kj} \right) (\, c_{ij}\,) \;=\; \left(\sum_\ell \left(\sum_k a_{ik}\; b_{k\ell} \right) c_{\ell j} \right) \\ &= \;\; \left(\sum_k a_{ik} \left(\sum_\ell b_{k\ell}\, c_{\ell j} \right) \right) \;=\; (\, a_{ij}\,) \left(\sum_\ell b_{i\ell}\, c_{\ell j} \right) \;=\; A\,(B\,C). \end{aligned}$$

These properties permit decomposition into blocks and exploitation of the many other aspects of matrices discussed below.
There is a square **zero matrix** $0 = (0)$, and an **identity matrix** $I = (\delta_{ij})$, where $\delta_{ij} = 1$ iff $i = j$, and 0 otherwise. All matrices A have a **additive inverse** $-A$, and some have a multiplicative inverse A^{-1}. Matrices also permit **scalar multiplication**:

$$\begin{array}{rclcll}
a\,B &=& C & \Leftrightarrow & a\,(b_{ij}) \;=\; (a\,b_{ij}) \;=\; (c_{ij}) & \text{scalar multiplication} \\
A+0 &=& A & \Leftrightarrow & (a_{ij}) \;+\; (0) \;=\; (a_{ij}) & \text{additive identity} \\
0+A &=& A & \Leftrightarrow & (0) \;+\; (a_{ij}) \;=\; (a_{ij}) & \\
A+(-A) &=& 0 & \Leftrightarrow & (a_{ij}) \;+\; (-a_{ij}) \;=\; (0) & \text{additive inverse} \\
(-A)+A &=& 0 & \Leftrightarrow & (-a_{ij}) \;+\; (a_{ij}) \;=\; (0) & \\
A\,I &=& A & \Leftrightarrow & (\sum_k a_{ik}\,\delta_{kj}) \;=\; (\sum_{k=j} a_{ik}) \;=\; (a_{ij}) & \text{multiplicative identity} \\
I\,A &=& A & \Leftrightarrow & (\sum_k \delta_{ik}\,a_{kj}) \;=\; (\sum_{i=k} a_{kj}) \;=\; (a_{ij}) & \\
A\,A^{-1} &=& I & \Leftrightarrow & (a_{ij})\;(a_{ij})^{-1} \;=\; (\delta_{ij}) & \text{multiplicative inverse} \\
A^{-1}\,A &=& I & \Leftrightarrow & (a_{ij})^{-1}\;(a_{ij}) \;=\; (\delta_{ij}) &
\end{array}$$

These properties define a **noncommutative ring** (matrix multiplication is not commutative).

0.2 Matrix Functions

0.2.1 Complex Conjugate

Matrix entries can be complex values $z = x + iy$, with **complex conjugate** $\overline{z} = x - iy$. The **matrix complex conjugate** of $A = (a_{ij})$ is $\overline{A} = (\overline{a_{ij}})$.

$$\begin{array}{rcl}
\overline{0} &=& 0 \\
\overline{I} &=& I \\
\overline{(\overline{A})} &=& A
\end{array}
\qquad
\begin{array}{rcl}
\overline{(A+B)} &=& \overline{A} + \overline{B} \\
\overline{(A+\overline{A})} &=& \overline{A} + A \\
\overline{(-A)} &=& -(\overline{A})
\end{array}
\qquad
\begin{array}{rcl}
\overline{(AB)} &=& \overline{A}\,\overline{B} \\
\overline{(ABC)} &=& \overline{A}\,\overline{B}\,\overline{C} \\
\overline{(A^{-1})} &=& (\overline{A})^{-1}
\end{array}$$

0.2.2 Transpose

The **ordinary transpose** of $A = (a_{ij})$ is $A^\top = (a_{ji})$; the **Hermitian transpose** is $A' = (\overline{a_{ji}})$. We will always use the Hermitian transpose.

$$\begin{array}{rcl}
0' &=& 0 \\
I' &=& I \\
(A')' &=& A
\end{array}
\qquad
\begin{array}{rcl}
(A+B)' &=& A' + B' \\
(A+A')' &=& A' + A \\
(-A)' &=& -(A')
\end{array}
\qquad
\begin{array}{rcl}
(AB)' &=& B'\;A' \\
(ABC)' &=& C'\;B'\;A' \\
(A')^{-1} &=& (A^{-1})'
\end{array}$$

Notice however that all of these properties hold if $'$ is replaced by $^\top$.

0.2.3 Trace

The **trace** of a $n \times n$ square matrix $A = (a_{ij})$ is $\text{tr}(A) = \sum_i a_{ii}$.

$$\begin{array}{rcl}
\text{tr}(0) &=& 0 \\
\text{tr}(I) &=& n
\end{array}
\qquad
\begin{array}{rcl}
\text{tr}(A+B) &=& \text{tr}(A) + \text{tr}(B) \\
\text{tr}(-A) &=& -\text{tr}(A)
\end{array}$$

0.2.4 Determinant

The **determinant** of a $n \times n$ square matrix $A = (a_{ij})$ is $\det(A) = \sum_{\sigma \in S_n} (-1)^\sigma \prod_i a_{i\,\sigma(i)}$ where S_n is the set of permutations σ on n items, $\sigma(i)$ is the value that σ permutes i to, and $(-1)^\sigma$ is the **sign of** σ, i.e., $+1$ or -1 depending on whether σ is an even permutation or odd permutation.

$$\begin{array}{rcl@{\qquad}rcl@{\qquad}rcl}
\det(0) & = & 0 & \det(\overline{A}) & = & \overline{\det(A)} & \det(AB) & = & \det(A)\ \det(B) \\
\det(I) & = & 1 & \det(A^\top) & = & \det(A) & \det(ABC) & = & \det(A)\ \det(B)\ \det(C) \\
\det(-A) & = & (-1)^n\ \det(A) & \det(A') & = & \overline{\det(A)} & \det(A^{-1}) & = & (\det(A))^{-1}
\end{array}$$

0.2.5 Matrix Power, Power Series, and Matrix Functions like *exp*

The p-**th power** of a square matrix A is the product of p copies of A: $A^p = \overbrace{A \cdots A}^{p \text{ copies}}$.
By convention, $A^0 = I$ and $A^{-p} = (A^{-1})^p$.
Matrix powers permit us to take functions defined by power series like $\exp(z) = \sum_{p \geq 0} z^p/p!$ and $1/(1-z) = 1 + z + z^2 + z^3 + \cdots$ and apply them to matrices:

$$\exp(A) = \sum_{p \geq 0} \frac{1}{p!} A^p, \qquad (I - A)^{-1} = \sum_{p \geq 0} A^p.$$

So, ignoring convergence issues, every complex-valued function $f(z) = \sum_p c_p z^p$ defined by a power series is also a square-matrix-valued function of square matrices $f(A) = \sum_p c_p A^p$.

0.3 Block Matrices and Submatrices

0.3.1 Block Matrix Product

If A and B are decomposed into block submatrices, we can define their product recursively:

$$A\,B = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}\,B_{11} + A_{12}\,B_{21} & A_{11}\,B_{12} + A_{12}\,B_{22} \\ A_{21}\,B_{11} + A_{22}\,B_{21} & A_{21}\,B_{12} + A_{22}\,B_{22} \end{pmatrix}.$$

For $n \times n$ square matrices A and B, this recursive property implies that the complexity of matrix multiplication follows the recurrence $T(n) = 8\,T(n/2)$, since there are 8 matrix products in the right-hand side. This has the solution $T(n) = O(n^3)$: multiplying large square matrices takes lots of time.

0.3.2 Kronecker Product / Tensor Product

If A and B have dimensions $m \times n$ and $p \times q$, their **tensor product** is the $mp \times nq$ matrix

$$A \otimes B = (\,a_{ij}\ B\,)$$

that has $i \times j$ blocks, each of size $p \times q$, where the ij-th block is the scalar product $a_{ij}\,B$.
Alternatively we can define $A \otimes B = C = (\,c_{ij}\,)$ where $c_{ij} = a_{i_1\,j_1}\,b_{i_2,j_2}$ and indices are defined by blocks: $i = (i_1 - 1)\,p + i_2$, $j = (j_1 - 1)\,q + j_2$, and $1 \leq i_2 \leq p$, $1 \leq j_2 \leq q$.

There is no identity or zero for this operation, but:

$$\begin{array}{rcll}
A \otimes (B + C) & = & A \otimes B + A \otimes C & \text{Distribution of tensor product over addition} \\
A \otimes (B \otimes C) & = & (A \otimes B) \otimes C & \text{Associativity of tensor product.}
\end{array}$$

The tensor product is useful for building large matrices having regular structure.

0.3.3 Hadamard Product

If A and B have the same dimensions, their **Hadamard product** is an element-wise product

$$A \cdot B \;=\; (\, a_{ij}\; b_{ij}\,).$$

The identity matrix $1 = (1)$ for this operation is a matrix whose entries are all 1. The Hadamard inverse matrix A^{-} for $A = (a_{ij})$ is clearly the matrix $A^{-} = (1/a_{ij})$ (which is defined if and only if all entries are nonzero). The Hadamard product also can be used to build matrices as a kind of multiplicative `filter'.

0.3.4 Direct Sum

If A has dimensions $m \times n$ and B has dimensions $p \times q$, their **direct sum** is the $(m+p) \times (n+q)$ matrix that has A along the upper left diagonal matrix, and B along the lower right diagonal:

$$A \oplus B \;=\; \left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array} \right).$$

The upper right $m \times q$ submatrix and lower left $p \times n$ submatrix have only zero entries.

0.3.5 Diagonals of Matrices

If A has dimensions $m \times n$, its **matrix diagonal** $\mathrm{diag}(A) = (a_{ii})$ --- a vector (single column matrix) of length $min(m, n)$.

0.4 Matrix Type Hierarchies

Figure 2 shows hierarchies for banded and nonnegative matrices, and Figure 3 shows a hierarchy of normal matrices.

0.4.1 Banded Matrices

Kind of matrix A	*condition on elements a_{ij} of A*
Diagonal	$a_{ij} = 0$ if $\lvert i-j\rvert > 0$
Tridiagonal	$a_{ij} = 0$ if $\lvert i-j\rvert > 1$
Upper Triangular	$a_{ij} = 0$ if $i-j > 0$
Lower Triangular	$a_{ij} = 0$ if $i-j < 0$
Upper Hessenberg	$a_{ij} = 0$ if $i-j > +1$
Lower Hessenberg	$a_{ij} = 0$ if $i-j < -1$
Toeplitz	$a_{ij} = f(i-j)$ for some real-valued function f

0.4.2 Nonnegative Matrices

Kind of matrix A	*condition on elements a_{ij} of A*
Nonnegative	$0 \le a_{ij}$
(Row) Stochastic	$0 \le a_{ij} \le 1$, rows total to 1 ($\sum_j a_{ij} = 1$ for $1 \le i \le n$)
Column Stochastic	$0 \le a_{ij} \le 1$, columns total to 1 ($\sum_i a_{ij} = 1$ for $1 \le j \le n$)
Doubly Stochastic	both Row Stochastic and Column Stochastic
Permutation	$a_{ij} \in \{0,\ 1\}$, Doubly Stochastic

Diagonal	$\begin{pmatrix} \times & & & & \\ & \times & & & \\ & & \times & & \\ & & & \times & \\ & & & & \times \end{pmatrix}$	Toeplitz	$\begin{pmatrix} e & f & g & h & i \\ d & e & f & g & h \\ c & d & e & f & g \\ b & c & d & e & f \\ a & b & c & d & e \end{pmatrix}$
Tridiagonal	$\begin{pmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \end{pmatrix}$	(Upper) Hessenberg	$\begin{pmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{pmatrix}$
Lower Triangular	$\begin{pmatrix} \times & & & & \\ \times & \times & & & \\ \times & \times & \times & & \\ \times & \times & \times & \times & \\ \times & \times & \times & \times & \times \end{pmatrix}$	Upper Triangular	$\begin{pmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{pmatrix}$

Figure 1: Banded Matrices

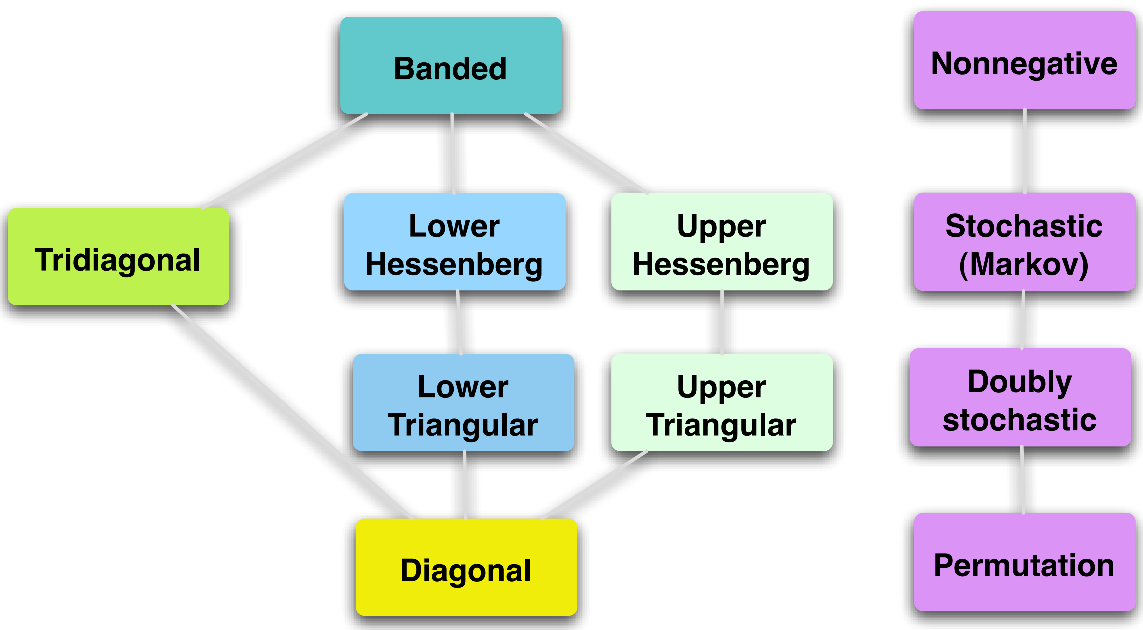

Figure 2: Hierarchies of Banded and Nonnegative matrices.

0.5 Vectors: Specialized Matrices

By convention, a **vector** is usually a $n \times 1$ (vertical) matrix, referred to as a **column vector**. The $1 \times n$ (horizontal) variant is a **row vector**.

$$\textit{column vector:}\ \ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \textit{row vector:}\ \ \mathbf{x}' = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}.$$

Thus matrix transpose converts between row and column vectors.

0.6 Scalar Products

The **scalar product** of two real n-vectors **x** and **y** is

$$\langle\, \mathbf{x},\, \mathbf{y}\,\rangle \;=\; \mathbf{x}'\ \mathbf{y} \;=\; \sum_{i=1}^{n} \overline{x_i}\, y_i .$$

The scalar product has many properties:

$$\begin{aligned}
\langle\, \mathbf{x},\, \mathbf{x}\,\rangle &\geq 0 \quad \text{(with equality iff } \mathbf{x} = \mathbf{0}\text{)} \\
\langle\, \mathbf{x},\, \mathbf{y}\,\rangle &= \overline{\langle\, \mathbf{y},\, \mathbf{x}\,\rangle} \\
\langle\, \lambda\, \mathbf{x},\, \mathbf{y}\,\rangle &= \lambda\, \langle\, \mathbf{x},\, \mathbf{y}\,\rangle \quad \text{if } \lambda \text{ is a nonnegative real value} \\
\langle\, \mathbf{x} + \mathbf{y},\, \mathbf{z}\,\rangle &= \langle\, \mathbf{x},\, \mathbf{z}\,\rangle + \langle\, \mathbf{y},\, \mathbf{z}\,\rangle.
\end{aligned}$$

Related terminology:

- If $\langle\, \mathbf{x},\, \mathbf{x}\,\rangle = 1$, **x** is a **unit vector**.
- If $\langle\, \mathbf{x},\, \mathbf{y}\,\rangle = 0$, we say **x** and **y** are **orthogonal**.

0.7 Vector Norms

Given a vector $\mathbf{x} = (\, x_1,\, x_2,\, \ldots,\, x_n\,)$, the **norm** of **x** is

$$||\mathbf{x}|| \;=\; \sqrt{\mathbf{x}'\ \mathbf{x}} \;=\; \sqrt{\langle\, \mathbf{x},\, \mathbf{x}\,\rangle} \;=\; \sqrt{\sum_i \overline{x_i}\, x_i} \;=\; \sqrt{\sum_i |x_i|^2}\,.$$

Some popular norms:

sum norm (Euclidean distance)	$\lVert\mathbf{x}\rVert_1$	$=$	$\sum_i \lvert x_i\rvert$
Euclidean norm (Manhattan distance)	$\lVert\mathbf{x}\rVert_2$	$=$	$\left(\sum_i \lvert x_i\rvert^2\right)^{1/2}$
max norm (Worst-case distance)	$\lVert\mathbf{x}\rVert_\infty$	$=$	$\max_i \lvert x_i\rvert$
general L^p norm (L^p distance)	$\lVert\mathbf{x}\rVert_p$	$=$	$(\sum_i \lvert x_i\rvert^p)^{1/p}$
A-**norm** (for positive definite matrices A)	$\lVert\mathbf{x}\rVert_A$	$=$	$(\langle\, \mathbf{x},\, A\,\mathbf{x}\,\rangle)^{1/2} = (\mathbf{x}'\ A\,\mathbf{x})^{1/2}$.

Notice that $\mathbf{x}/||\mathbf{x}||_2$ is a unit vector (when $\mathbf{x} \neq \mathbf{0}$).

Given a norm we can define a corresponding **distance** measure:

$$d(\mathbf{x},\mathbf{y}) \quad = \quad ||\mathbf{x}-\mathbf{y}|| \quad = \quad \sqrt{\langle\, \mathbf{x}-\mathbf{y},\, \mathbf{x}-\mathbf{y}\,\rangle}.$$

This distance measure is a **metric**:

- nonnegative: $d(\mathbf{x},\mathbf{y}) \geq 0$ (and is zero iff $\mathbf{x} = \mathbf{y}$)
- symmetric: $d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$.
- triangle inequality: $d(\mathbf{x},\mathbf{z}) \leq d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})$.

0.8 Projection of Vectors on Vectors

Given two vectors **u** and **v**, the **projection** of $\mathbf{v}$ on $\mathbf{u}$ $= \dfrac{\langle\, \mathbf{u},\, \mathbf{v}\,\rangle}{\langle\, \mathbf{u},\, \mathbf{u}\,\rangle}\,\mathbf{u} = \dfrac{\langle\, \mathbf{u},\, \mathbf{v}\,\rangle}{||\mathbf{u}||_2^2}\,\mathbf{u} = \langle\, \mathbf{e},\, \mathbf{v}\,\rangle\, \mathbf{e}$
where $\mathbf{e} = \mathbf{u}/||\mathbf{u}||_2$ is a unit vector. If **u** and **v** are orthogonal, the result is a zero vector.

0.9 Bases, Coordinate Systems, Vector Spaces

A **basis** is a set of vectors

$$B \;=\; \{\, \mathbf{e}_1,\; \ldots,\; \mathbf{e}_n \,\}$$

that can be used to build other vectors **v** as **linear combinations**

$$\mathbf{v} \;=\; v_1\, \mathbf{e}_1 \;+\; \cdots \;+\; v_n\, \mathbf{e}_n$$

with numeric coefficients v_i. Any orthonormal basis $B = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ defines a n-**dimensional coordinate system**: each such vector **v** can then be represented by the **coordinates**

$$\mathbf{v} \;\Leftrightarrow\; (\, v_1,\, \ldots\, v_n\,).$$

The **vector space** defined by a basis B over a set of values R (usually the real numbers) is the set of all possible linear combinations of B's basis vectors:

$$\text{vector space defined by } B \;=\; \left\{\; \mathbf{x} \;\middle|\; \mathbf{x} \;=\; \sum_{i=1}^{n} v_i\, \mathbf{e}_i, \quad \text{each } v_i \text{ is a value in } R \;\right\}.$$

0.10 Cross Product

For real 3D vectors **u** and **v**, the **vector cross product** is

$$\mathbf{u} \times \mathbf{v} \;=\; \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \times \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \;=\; \begin{pmatrix} (u_2 v_3 - u_3 v_2) \\ (-u_1 v_3 + u_3 v_1) \\ (u_1 v_2 - u_2 v_1) \end{pmatrix}.$$

This definition satisfies $(\mathbf{v} \times \mathbf{u}) = -(\mathbf{u} \times \mathbf{v})$. Also the result is orthogonal to **u**:

$$\begin{aligned}
\langle\, \mathbf{u},\, \mathbf{u} \times \mathbf{v}\,\rangle \;&=\; \left\langle \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \begin{pmatrix} (u_2 v_3 - u_3 v_2) \\ (-u_1 v_3 + u_3 v_1) \\ (u_1 v_2 - u_2 v_1) \end{pmatrix} \right\rangle \\
&=\; u_1\,(u_2 v_3 - u_3 v_2) \;+\; u_2\,(-u_1 v_3 + u_3 v_1) \;+\; u_3\,(u_1 v_2 - u_2 v_1) \\
&=\; u_1 u_2 v_3 - u_1 u_3 v_2 - u_2 u_1 v_3 + u_2 u_3 v_1 + u_3 u_1 v_2 - u_3 u_2 v_1 \\
&=\; (u_1 u_2 v_3 - u_2 u_1 v_3) + (u_3 u_1 v_2 - u_1 u_3 v_2) + (u_2 u_3 v_1 - u_3 u_2 v_1) \\
&=\; 0.
\end{aligned}$$

Because $(\mathbf{v} \times \mathbf{u}) = -(\mathbf{u} \times \mathbf{v})$, the result is also orthogonal to **v**.

0.11 Linear Transforms

A function f is called **linear** if

$$\begin{aligned} f(x + y) &= f(x) + f(y) \\ f(a\,x) &= a\,f(x) \end{aligned}$$

where a is a scalar. In particular, any matrix transformation of vectors is linear:

$$f(\mathbf{x}) = A\,\mathbf{x}$$

For all discussions of eigenvalues below A is always assumed to be a real symmetric matrix.
A **diagonal matrix** is a matrix whose only nonzero entries are on the diagonal:

$$D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

When the values λ_i are real values, the matrix D is called a **dilation**.
An **orthogonal matrix** is a matrix whose columns define an **orthonormal basis**:

$$Q = \left(\mathbf{e}_1 \,\middle|\, \mathbf{e}_2 \,\middle|\, \cdots \,\middle|\, \mathbf{e}_n \right).$$

Every orthogonal matrix satisfies:

$$Q'\,Q = Q\,Q' = I.$$

Theorem 1 *If A is a real symmetric matrix, then we can find a real diagonal matrix L and an orthogonal matrix Q such that A can be decomposed into the product*

$$A = Q\,L\,Q'.$$

Furthermore, if Q and L have the form

$$Q = \left(\mathbf{e}_1 \,\middle|\, \mathbf{e}_2 \,\middle|\, \cdots \,\middle|\, \mathbf{e}_n \right) \qquad L = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

then the orthonormal columns $\mathbf{e}_j$ of Q are eigenvectors of A, and the diagonal elements λ_i of L are eigenvalues of A.

In the 1820s Cauchy studied the eigenvalues of real symmetric matrices, i.e., symmetric linear systems of the form

$$\begin{aligned} a\,x + b\,y &= \lambda\,x \\ b\,x + d\,y &= \lambda\,y \end{aligned} \quad \Longleftrightarrow \quad \begin{pmatrix} a & b \\ b & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix}.$$

If $b \neq 0$, this system can be reduced to a quadratic equation, and therefore there are two eigenvalue solutions $\lambda = ((a+d) \pm \sqrt{(a-d)^2 + 4\,b^2})\,/\,2$.

0.12 Determinants, Rank, and Matrix Norms

The determinant of a $n \times n$ matrix A is the product of its eigenvalues:

$$det(A) \;=\; \lambda_1(A) \,\cdots\, \lambda_n(A).$$

The **rank** of a matrix A is its number of nonzero eigenvalues.

A **matrix norm** is a measure of *the maximum dilation of a matrix*:

$$||A|| \;=\; \max_{||\mathbf{v}||=1} \; ||\, A\, \mathbf{v}\, ||.$$

Since the maximum dilation is the maximum eigenvalue of A, $||A||$ is the maximum eigenvalue of A: $||A|| = |\lambda_{\max}|$. Matrix norms have following properties:

1. $||A|| \geq 0$.
2. $||A|| = 0$ if and only if $A = 0$.
3. $||cA|| = |c|||A||$ for all complex scalars c.
4. $||A + B|| \leq ||A|| + ||B||$.
5. $||AB|| \leq ||A||\,||B||$.

By verifying these properties, we can establish that the following commonly-encountered measures are, in fact, matrix norms:

Spectral norm	$\|\|A\|\|_2$	$=$	$\|\lambda_{\max}\|$
Infinity (max row sum) norm	$\|\|A\|\|_\infty$	$=$	$\max_i \sum_j \|a_{ij}\|$
L_1 **(max col sum) norm**	$\|\|A\|\|_1$	$=$	$\max_j \sum_i \|a_{ij}\|$
Frobenius norm	$\|\|A\|\|_F$	$=$	$\sqrt{\sum_{i,j} a_{ij}^2}.$

The **spectral radius** of A is $|\lambda_{\max}| = ||A||_2$, its largest eigenvalue:

$$||A||_2 \;=\; \max_{||\mathbf{x}||} \; ||A\mathbf{x}||_2 \,/\, ||\mathbf{x}||_2 \;=\; \max_{||\mathbf{e}||=1} \; ||A\mathbf{e}||_2 \;=\; |\lambda_{max}(A)|.$$

0.13 Using Eigenstructure to Invert a Matrix

Given the eigendecomposition $A \;=\; Q\, L\, Q'$,

$$A^{-1} \;=\; (Q\, L\, Q'\,)^{-1} \;=\; (Q'\,)^{-1}\, L^{-1}\, (Q)^{-1} \;=\; Q\, L^{-1}\, Q'\, .$$

Thus A is **invertible** (**nonsingular**) if and only if it has no zero eigenvalues.

0.14 Characteristic Polynomials

The **characteristic polynomial of** A is defined as

$$P_A(\lambda) \;=\; \det(A \;-\; \lambda\, I)$$

where λ is a variable and I is the identity matrix.
An **eigenvalue** of A is then any root of this polynomial. A corresponding **eigenvector v** is any vector satisfying $A\mathbf{v} \;=\; \lambda\mathbf{v}$, or equivalently: $(A \;-\; \lambda I)\,\mathbf{v} \;=\; \mathbf{0}$. Thus $(A \;-\; \lambda I)$ cannot have a matrix inverse, and its determinant must be zero.

0.14.1 Normal Matrices

Kind of matrix A	*condition on elements a_{ij} of A*
Normal	$A\ A' = A'\ A$
Unitary	$A' = A^{-1}$
Orthogonal	$A' = A^{-1}$, A is real
Rotation	$A' = A^{-1}$, A is real, $\det(A) = +1$
Reflection	$A' = A^{-1}$, A is real, $\det(A) = -1$
Hermitian	$A' = A$
Symmetric	$A' = A$, A is real
Positive Definite	$A' = A$, A is real, A has only positive eigenvalues
Negative Definite	$A' = A$, A is real, A has only negative eigenvalues
Skew-Hermitian	$A' = -A$
Skew-Symmetric	$A' = -A$, A is real

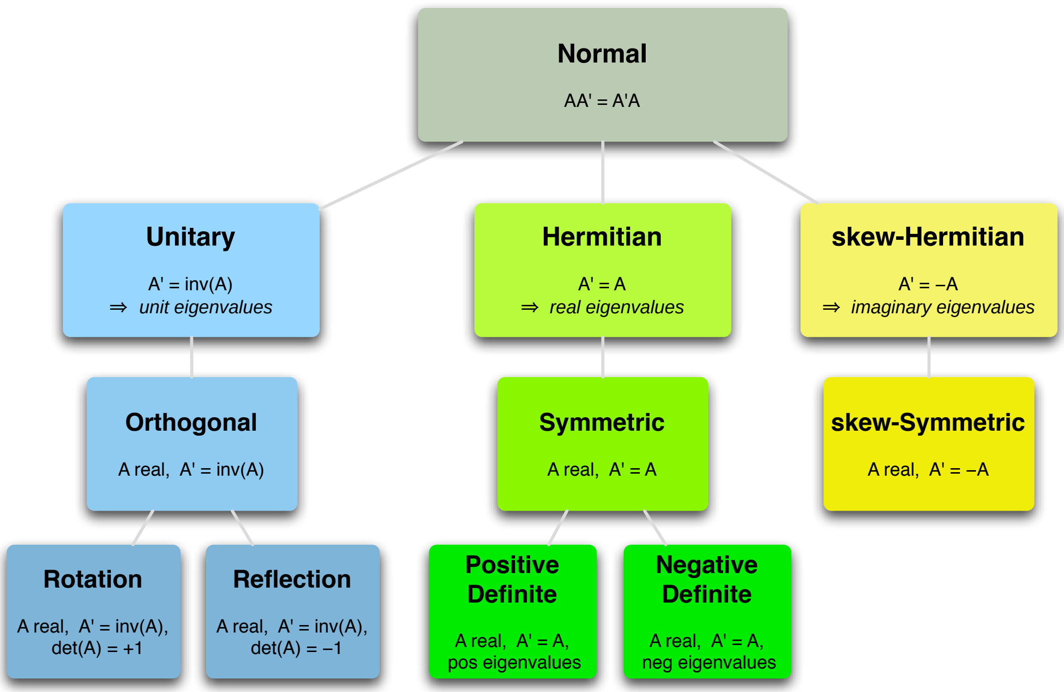

Figure 3: Hierarchy of Normal matrices.