Introduction to Data Assimilation

Some useful linear algebra

Vector and matrix notation

$\mathbf{x} \in \Re^N$ denotes a column vector with *N* (real) entries, i.e: $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$.

x_n is the element in the n^{th} position of the vector $\mathbf{x}$.

$\mathbf{x}^T$ denotes the transpose of $\mathbf{x}$; it becomes a row vector: $\mathbf{x}^T = [x_1 \quad x_2 \quad \cdots \quad x_N]$.

$\mathbf{A} \in \Re^{N \times M}$ denotes a matrix with *N* rows and *M* columns: $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1M} \\ a_{21} & a_{22} & \cdots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NM} \end{bmatrix}$.

a_{nm} is the element of $\mathbf{A}$ located in the n^{th} row and m^{th} column.

We can consider each column of a matrix to be a vector, then we can express it as:

$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_M]$, where the m^{th} column is $\mathbf{a}_m = \begin{bmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{Nm} \end{bmatrix}$.

The column rank of a matrix is the number of linearly independent columns; the row rank of a matrix is the number of linearly independent rows. Both ranks coincide. For a matrix $\mathbf{A} \in \Re^{N \times M}$, $rank(\mathbf{A}) \leq \min(N, M)$. A matrix is said to be full rank if its rank has the maximum possible value. Otherwise, it is called rank deficient.

$\mathbf{A}^T$ denotes the transpose of $\mathbf{A}$, it makes rows into columns and columns into rows. E.g., for $N = 3$ and $M = 2$:

$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$, then $\mathbf{A}^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \end{bmatrix}$

A square matrix is one that has the same number of rows and columns, i.e. $M = N$.

A symmetric matrix is a square matrix that is equal to its transpose, i.e. $\mathbf{A} = \mathbf{A}^T$.

The diagonal of a square matrix is the collection of values a_{nm} where $n = m$, i.e. $\{a_{11}, a_{22}, \cdots, a_{NN}\}$.

The trace of a matrix is the sum of the values in the diagonal, i.e. $trace(\mathbf{A}) = \sum_{n=1}^{N} a_{nn}$

A diagonal (square) matrix has zero values everywhere except on the diagonal:

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{NN} \end{bmatrix}.$$

$\mathbf{I}$ denotes the identity matrix; it is a diagonal matrix with ones on the diagonal:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Basic vector and matrix operations

The product of a scalar λ and a matrix $\mathbf{A} \in \Re^{N \times M}$ is: $\lambda \mathbf{A} = \begin{bmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1M} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{N1} & \lambda a_{N2} & \cdots & \lambda a_{NM} \end{bmatrix}$.

For two matrices $\mathbf{A} \in \Re^{N \times M}$ and $\mathbf{B} \in \Re^{M \times L}$, let the matrix product be $\mathbf{C} = \mathbf{AB}$. Then $\mathbf{C} \in \Re^{N \times L}$ and $c_{nl} = \sum_{m=1}^{M} a_{nm} b_{ml}$. Note that the operation is defined only if the number of columns of the left factor is equal to the number of rows of the right factor.

For two matrices $\mathbf{A} \in \Re^{N \times M}$ and $\mathbf{B} \in \Re^{M \times N}$, both products $\mathbf{C} = \mathbf{AB}$ and $\mathbf{D} = \mathbf{BA}$ exist. However, in general $\mathbf{C} \neq \mathbf{D}$, i.e. matrix multiplication is not commutative.

The transpose of a product is the product of the transposes of the factors in reverse order, i.e. $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$.

The inverse of a square matrix $\mathbf{A} \in \Re^{N\times N}$ is denoted as $\mathbf{A}^{-1} \in \Re^{N\times N}$, and satisfies the relationship $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. The inverse of a matrix may or may not exist (depending on the characteristics of $\mathbf{A}$), and finding it is not a trivial exercise.

The inverse of a diagonal matrix $\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{NN} \end{bmatrix}$ is simply $\mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{a_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{a_{NN}} \end{bmatrix}$.

The inverse of a product is $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, given that both $\mathbf{A}^{-1}$ and $\mathbf{B}^{-1}$ exist.

The Schur (or Hadamard) product of two matrices $\mathbf{A} \in \Re^{N\times M}$ and $\mathbf{B} \in \Re^{N\times M}$ is an entrywise product, and is denoted as $\mathbf{C} = \mathbf{A} \circ \mathbf{B}$. In this case, $\mathbf{C} \in \Re^{N\times M}$ and $c_{nm} = a_{nm}b_{nm}$. This product is defined for any two matrices with the same dimensions, and is commutative, i.e. $\mathbf{A} \circ \mathbf{B} = \mathbf{B} \circ \mathbf{A}$.

The inner product (or dot product) of two vectors $\mathbf{x} \in \Re^{N}$ and $\mathbf{z} \in \Re^{N}$ is denoted as $\mathbf{x} \cdot \mathbf{z}$, $\langle \mathbf{x}, \mathbf{z} \rangle$ and defined as $u = \mathbf{x}^T\mathbf{z} = \mathbf{z}^T\mathbf{x} = \sum_{n=1}^{N} x_n z_n$. Note that the result is a scalar.

In particular, the inner product of a vector with itself is $\mathbf{x}^T\mathbf{x} = \sum_{n=1}^{N} x_n^2$, the square of the Euclidean norm of $\mathbf{x}$.

The outer product of two vectors $\mathbf{x} \in \Re^{N}$ and $\mathbf{z} \in \Re^{M}$ is defined as $\mathbf{U} = \mathbf{x}\mathbf{z}^T$. The result is a matrix $\mathbf{U} \in \Re^{NM}$, where $u_{nm} = a_n b_m$. Note that the outer product is not commutative and in fact $\mathbf{x}\mathbf{z}^T$ and $\mathbf{z}\mathbf{x}^T$ have different dimensions.