Data Assimilation Background

Elias D. Nino-Ruiz, Ph.D.

Universidad del Norte

enino@uninorte.edu.co

February 2, 2022

Outline

Vectors and Matrices I

- Vectors are always column vectors,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n \times 1}$$

  x_i denotes the i-th element of $\mathbf{x}$.

- The transpose of a vector reads,

$$\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \in \mathbb{R}^{1 \times n}.$$

Vectors and Matrices II

- The inner product of two vectors $\mathbf{x} \in \mathbb{R}^{n \times 1}$ and $\mathbf{y} \in \mathbb{R}^{n \times 1}$ reads,

$$\mathbf{x}^T \cdot \mathbf{y} = \sum_{i=1}^{n} x_i \cdot y_i \in \mathbb{R}\,,$$

  where y_i is the i-th component of vector $\mathbf{y}$.

- The outer product of two vectors $\mathbf{x} \in \mathbb{R}^{n \times 1}$ and $\mathbf{z} \in \mathbb{R}^{m \times 1}$ is given by,

$$\mathbf{x} \cdot \mathbf{z}^T = \begin{bmatrix} x_1 \cdot z_1 & x_1 \cdot z_2 & \dots & x_1 \cdot z_m \\ x_2 \cdot z_1 & x_2 \cdot z_2 & \dots & x_2 \cdot z_m \\ \vdots & \vdots & \ddots & \vdots \\ x_n \cdot z_1 & x_n \cdot z_2 & \dots & x_n \cdot z_m \end{bmatrix} \in \mathbb{R}^{n \times m}\,,$$

  where z_i is the i-th element of vector $\mathbf{z}$.

Vectors and Matrices III

- Square matrices are of the form,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

- General matrices are of the form,

$$\mathbf{B} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \in \mathbb{R}^{m \times n}$$

The components of **A** can be referenced as $\{\mathbf{A}\}_{i,j}$, for $1 \leq i \leq m$ and $1 \leq j \leq n$.

- The transpose of a matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$,

$$\mathbf{B}^T = \left\{\mathbf{B}^T\right\}_{i,j} = \{\mathbf{B}\}_{j,i} \in \mathbb{R}^{n \times m}$$

- The matrix $\mathbf{B}$ is said to be square if $m = n$.
- A square is positive definite if there is a positive scalar α such that,

$$\mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x} \geq \alpha \cdot \mathbf{x}^T \cdot \mathbf{x}\,,\ \text{for all } \mathbf{x} \in \mathbb{R}^{n \times 1}\,,$$

  or

$$\mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x} > 0,\ \text{for all } \mathbf{x} \in \mathbb{R}^{n \times 1}\text{,except } \mathbf{x} = \mathbf{0}\,.$$

- A square is positive semidefinite if

$$\mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x} \geq 0,\ \text{for some } \mathbf{x} \in \mathbb{R}^{n \times 1},\ \text{except } \mathbf{x} = \mathbf{0}\ .$$

Vectors and Matrices V

- The diagonal of the matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$ consists of elements $\{A\}_{i,i}$, $1 \leq i \leq \min(n, m)$.
- The matrix $\mathbf{A}$ is lower triangular if $\{\mathbf{A}\}_{i,j} = 0$ whenever $i < j$.
- The matrix $\mathbf{A}$ is upper triangular if $\{\mathbf{A}\}_{i,j} = 0$ whenever $i > j$.
- A square matrix is non-singular if for any vector $\mathbf{b} \in \mathbb{R}^{n \times 1}$, there exists $\mathbf{x} \in \mathbb{R}^{n \times 1}$ such that,

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}\,.$$

- The identity matrix, denoted by $\mathbf{I}$, is the square diagonal matrix whose diagonal elements are all 1.
- For non-singular matrices $\mathbf{A} \in \mathbb{R}^{n \times n}$, there is a unique $n \times n$ matrix $\mathbf{C} \in \mathbb{R}^{n \times n}$ such that,

$$\mathbf{A} \cdot \mathbf{C} = \mathbf{C} \cdot \mathbf{A} = \mathbf{I}\,,$$

we denote $\mathbf{C} = \mathbf{A}^{-1}$ and call it the inverse of $\mathbf{A}$.

Vectors and Matrices VI

- A square matrix $\mathbf{Q}$ is orthogonal if

$$\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{Q} \cdot \mathbf{Q}^T = \mathbf{I}$$

Norms I

- For a vector $\mathbf{x} \in \mathbb{R}^{n\times 1}$, we define the following norms,

$$\begin{aligned} \|\mathbf{x}\|_1 &= \sum_{i=1}^{n} |x_i| \\ \|\mathbf{x}\|_2 &= \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \\ \|\mathbf{x}\|_\infty &= \max_{1\le i\le n} |x_i| \end{aligned}$$

All these norms measure the length pf the vector in some sense, and they are equivalent in the sense that each one is bounded above and below by a multiple of the other, for instance,

$$\|\mathbf{x}\|_\infty \le \|x\|_2 \le \sqrt{n}\cdot\|\mathbf{x}\|_\infty \,, \|\mathbf{x}\|_\infty \le \|x\|_1 \le n\cdot\|\mathbf{x}\|_\infty$$

Norms II

- Some properties,

$$
\begin{aligned}
\|\mathbf{x}+\mathbf{z}\| &\leq \|\mathbf{x}\|+\|z\| \\
\|\mathbf{x}\| &= 0 \Rightarrow \mathbf{x}=\mathbf{0} \\
\|\alpha \cdot \mathbf{x}\| &= |\alpha| \cdot \|\mathbf{x}\|
\end{aligned}
$$

- Another interesting property that holds for $\|.\|_2$ is the Cauchy-Schwarz inequality,

$$
|\mathbf{x}^T \cdot \mathbf{y}| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\| \, .
$$

Norms III

- For a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$, matrix norms are defined as follows,

$$
\begin{aligned}
\|\mathbf{A}\|_1 &= \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}| \\
\|\mathbf{A}\|_2 &= \sigma_{\max} \left(\mathbf{A}^T \cdot \mathbf{A} \right)^{1/2} \\
\|\mathbf{A}\|_\infty &= \max_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|
\end{aligned}
$$

- For the Euclidean norm $\|.\|_2$, the following property holds,

$$\|\mathbf{A} \cdot \mathbf{B}\| \le \|\mathbf{A}\| \cdot \|\mathbf{B}\| \ ,$$

for all matrices **A** and **B** with consistent dimensions.

Norms IV

- The condition number of a non-singular matrix is defined as

$$\kappa\left(\mathbf{A}\right) = \left\|\mathbf{A}\right\| \cdot \left\|\mathbf{A}^{-1}\right\| \,.$$

- The Frobenius norm $\left\|\mathbf{A}\right\|_F$ of the matrix $\mathbf{A}$ is defined as follows,

$$\left\|\mathbf{A}\right\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}$$

Subspaces

- Given the Euclidean space $\mathbb{R}^n$, the subset $\mathcal{S} \subset \mathbb{R}^n$ is a sub-space of $\mathbb{R}^n$ is the following property holds,

$$\alpha \cdot x + \beta \cdot y \in \mathcal{S}, \text{ for all } \alpha,\, \beta \in \mathbb{R}.$$

- A set of vectors

$$\{\mathbf{x}_1,\, \mathbf{x}_2,\, \ldots,\, \mathbf{x}_m\} \in \mathbb{R}^{n \times 1}\,,$$

is called linearly independent set if there are no real numbers $\{\alpha_i\}_{i=1}^m$ such that,

$$\sum_{i=1}^{m} \alpha_i \cdot \mathbf{x}_i = \mathbf{0}\,,$$

unless we make the trivial choice $\alpha_1 = \alpha_2 = \ldots = \alpha_m = 0$.

- Spanning set,

$$\mathbf{x} = \sum_{i=1}^{m} \alpha_i \cdot \mathbf{x}_i, \text{ for } \mathbf{x} \in \mathcal{S}\,.$$

Eigenvalues, Eigenvectors and the Singular Value Decomposition I

- A scalar value λ is an eigenvalue of the $n \times n$ matrix $\mathbf{A}$ if there is a nonzero vector $\mathbf{q} \in \mathbb{R}^{n\times 1}$,

$$\mathbf{A} \cdot \mathbf{q} = \lambda \cdot \mathbf{q} \,,$$

  the matrix $\mathbf{A}$ is nonsingular if none of its eigenvalues are zero.
- The eigen values of symmetric matrices are all real numbers, while nonsymmetric matrices may have imaginary eigenvalues.

Eigenvalues, Eigenvectors and the Singular Value Decomposition II

- All matrices $\mathbf{A}$ (not necessarily square) can be decomposed as a product of three matrices with special properties. When $\mathbf{A} \in \mathbb{R}^{m \times n}$ with $m > n$, its Singular Value Decomposition has the form,

$$\mathbf{A} = \mathbf{U} \cdot \begin{bmatrix} \mathbf{S} \\ \mathbf{0} \end{bmatrix} \cdot \mathbf{V}^T \,,$$

  where $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ are orthogonal matrices, and $\mathbf{S} \in \mathbb{R}^{n \times n}$ is a diagonal matrix whose diagonal elements are the singular values of $\mathbf{A}$ in descending order,

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \,. \quad (1)$$

Eigenvalues, Eigenvectors and the Singular Value Decomposition III

- For $n \geq m$

$$\mathbf{A} = \mathbf{U} \cdot \begin{bmatrix} \mathbf{S} & \mathbf{0} \end{bmatrix} \cdot \mathbf{V}^T \,,$$

where $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ are orthogonal matrices, and $\mathbf{S} \in \mathbb{R}^{m \times m}$ is a diagonal matrix whose diagonal elements are the singular values of $\mathbf{A}$ in descending order,

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \,.$$

Eigenvalues, Eigenvectors and the Singular Value Decomposition IV

- Spectral decomposition of $\mathbf{A} \in \mathbb{R}^{n \times n}$

$$\mathbf{A} = \sum_{i=1}^{n} \sigma_i \cdot \mathbf{u}_i \cdot \mathbf{v}_i^T \,,$$

where

$$\begin{aligned}
\mathbf{S} &= \mathbf{diag}\left\{\sigma_1,\, \sigma_2,\, \ldots,\, \sigma_n\right\} \in \mathbb{R}^{n \times n} \\
\mathbf{U} &= \left[\mathbf{u}_1,\, \mathbf{u}_2,\, \ldots,\, \mathbf{u}_n\right] \in \mathbb{R}^{n \times n}\,, \\
\mathbf{V} &= \left[\mathbf{v}_1,\, \mathbf{v}_2,\, \ldots,\, \mathbf{v}_n\right] \in \mathbb{R}^{n \times n}\,,
\end{aligned}$$

from which, functional over matrices are defined,

$$f(\mathbf{A}) = \sum_{i=1}^{n} f\left(\sigma_i\right) \cdot \mathbf{u}_i \cdot \mathbf{v}_i^T \,,$$

Eigenvalues, Eigenvectors and the Singular Value Decomposition V

- When $\mathbf{A} \in \mathbb{R}^{n \times n}$ is symmetric,

$$\mathbf{A} = \sum_{i=1}^{n} \sigma_i \cdot \mathbf{u}_i \cdot \mathbf{v}_i^T = \sum_{i=1}^{n} \sigma_i \cdot \mathbf{u}_i \cdot \mathbf{u}_i^T = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{U}^T \,. \tag{2}$$

Determinant and Trace I

- The trace of $\mathbf{A} \in \mathbb{R}^{n \times n}$ is defined by,

$$\mathrm{trace}\left(\mathbf{A}\right) = \sum_{i=1}^{n} \left\{\mathbf{A}\right\}_{i,i} \,,$$

  you can show that,

$$\mathrm{trace}\left(\mathbf{A}\right) = \sum_{i=1}^{n} \sigma_i \,,$$

  where σ_i denotes the i-th singular value of $\mathbf{A}$.

- The determinant of $\mathbf{A} \in \mathbb{R}^{n \times n}$ reads,

$$\det\left(\mathbf{A}\right) = \prod_{i=1}^{n} \sigma_i \,,$$

  where σ_i denotes the i-th singular value of $\mathbf{A}$.

Determinant and Trace II

- The determinant has several revealing properties,
  1. $\det(\mathbf{A}) = 0$ if and only if $\mathbf{A}$ is singular.
  2. $\det(\mathbf{A}) \cdot \det(\mathbf{B}) = \det(\mathbf{A} \cdot \mathbf{B})$.
  3. $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.
- Please recall,
  1. Gaussian elimination.
  2. LU factorization.
  3. Cholesky factorization.