

MATHS PRIMER FOR COURSE ON DA METHODS
(NERC/NCEO/DARC/UOR)

RNB, Univ. of Reading, NCEO, June 2025

1. VECTORS AND MATRICES

1.1. **Vector representation of information (n elements).**

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix}, \quad v_i = (\mathbf{v})_i.$$

1.2. **Matrix ($m \times n$ elements).**

$$\mathbf{N} = \begin{pmatrix} N_{11} & \cdots & N_{1j} & \cdots & N_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ N_{i1} & \cdots & N_{ij} & \cdots & N_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ N_{m1} & \cdots & N_{mj} & \cdots & N_{mn} \end{pmatrix}, \quad N_{ij} = (\mathbf{N})_{ij}.$$

1.3. **Matrix ($m \times n$) acting on a vector.**

$$\mathbf{v}^b = \mathbf{N}\mathbf{v}^a, \quad v_i^b = \sum_{j=1}^n N_{ij} v_j^a, \quad 1 \leq i \leq m \text{ as below:}$$

$$\begin{pmatrix} v_1^b \\ \vdots \\ v_i^b \\ \vdots \\ v_m^b \end{pmatrix} = \begin{pmatrix} N_{11} & \cdots & N_{1j} & \cdots & N_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ N_{i1} & \cdots & N_{ij} & \cdots & N_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ N_{m1} & \cdots & N_{mj} & \cdots & N_{mn} \end{pmatrix} \begin{pmatrix} v_1^a \\ \vdots \\ v_j^a \\ \vdots \\ v_n^a \end{pmatrix} = \begin{pmatrix} N_{11}v_1^a + \cdots + N_{1j}v_j^a + \cdots + N_{1n}v_n^a \\ \vdots \\ N_{i1}v_1^a + \cdots + N_{ij}v_j^a + \cdots + N_{in}v_n^a \\ \vdots \\ N_{m1}v_1^a + \cdots + N_{mj}v_j^a + \cdots + N_{mn}v_n^a \end{pmatrix}.$$

1.4. **Identity/unit matrix ($p \times p$ elements).**

$$\mathbf{I}_p = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \quad (\mathbf{I}_p)_{ij} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

1.5. **Matrix addition ($m \times n$ elements).**

$$\mathbf{N} = \mathbf{N}^a + \mathbf{N}^b, \quad N_{ij} = N_{ij}^a + N_{ij}^b \text{ as below:}$$

$$\begin{pmatrix} N_{11}^a & \cdots & N_{1n}^a \\ \vdots & \vdots & \vdots \\ N_{m1}^a & \cdots & N_{mn}^a \end{pmatrix} + \begin{pmatrix} N_{11}^b & \cdots & N_{1n}^b \\ \vdots & \ddots & \vdots \\ N_{m1}^b & \cdots & N_{mn}^b \end{pmatrix} = \begin{pmatrix} N_{11}^a + N_{11}^b & \cdots & N_{1n}^a + N_{1n}^b \\ \vdots & \ddots & \vdots \\ N_{m1}^a + N_{m1}^b & \cdots & N_{mn}^a + N_{mn}^b \end{pmatrix}.$$

1

1.6. Matrix multiplication ($m \times n$ elements - an $m \times p$ matrix multiplied by an $p \times n$ matrix).

$$\mathbf{N} = \mathbf{N}^a \mathbf{N}^b, \quad N_{ij} = \sum_{k=1}^p N_{ik}^a N_{kj}^b \text{ as below:}$$

$$\begin{pmatrix} N_{11}^a & \cdots & N_{1p}^a \\ \vdots & \ddots & \vdots \\ N_{m1}^a & \cdots & N_{mp}^a \end{pmatrix} \begin{pmatrix} N_{11}^b & \cdots & N_{1n}^b \\ \vdots & \ddots & \vdots \\ N_{p1}^b & \cdots & N_{pn}^b \end{pmatrix} = \begin{pmatrix} N_{11}^a N_{11}^b + \cdots + N_{1p}^a N_{p1}^b & \cdots & N_{11}^a N_{1n}^b + \cdots + N_{1p}^a N_{pn}^b \\ \vdots & \ddots & \vdots \\ N_{m1}^a N_{11}^b + \cdots + N_{mp}^a N_{p1}^b & \cdots & N_{m1}^a N_{1n}^b + \cdots + N_{mp}^a N_{pn}^b \end{pmatrix}.$$

In general, matrices are non-commutative $\mathbf{N}^a \mathbf{N}^b \neq \mathbf{N}^b \mathbf{N}^a$.

1.7. Matrix transpose ($\mathbf{N}^a$ has $m \times n$ elements, $\mathbf{N}^b$ has $n \times m$ elements).

$$\text{If } \mathbf{N}^b = \mathbf{N}^{aT}, \quad N_{ij}^b = N_{ji}^a, \text{ e.g.:}$$

$$\mathbf{N}^a = \begin{pmatrix} N_{11}^a & N_{12}^a & N_{13}^a \\ N_{21}^a & N_{22}^a & N_{23}^a \end{pmatrix}, \quad \mathbf{N}^b = \begin{pmatrix} N_{11}^a & N_{21}^a \\ N_{12}^a & N_{22}^a \\ N_{13}^a & N_{23}^a \end{pmatrix}.$$

If $\mathbf{N}^a = \mathbf{N}^{aT}$ then matrix $\mathbf{N}^a$ is symmetric (only square matrices ($m = n$) can be symmetric).

1.8. Transpose of a product of matrices.

$$(\mathbf{N}^a \mathbf{N}^b)^T = \mathbf{N}^{bT} \mathbf{N}^{aT}.$$

1.9. Matrix inversion ($n \times n$ elements). Let $\mathbf{N}$ be a square ($n \times n$) non-singular matrix.

$$\text{If } \mathbf{v}^b = \mathbf{N} \mathbf{v}^a, \text{ then } \mathbf{v}^a = \mathbf{N}^{-1} \mathbf{v}^b.$$

$$\text{In general } (\mathbf{N}^{-1})_{ij} \neq (\mathbf{N})_{ij}^{-1}.$$

$$\text{For } n = 2, \quad \mathbf{N} = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix}, \quad \mathbf{N}^{-1} = \frac{1}{\det(\mathbf{N})} \begin{pmatrix} N_{22} & -N_{12} \\ -N_{21} & N_{11} \end{pmatrix}, \quad \det(\mathbf{N}) = N_{11}N_{22} - N_{12}N_{21}.$$

If $\mathbf{N}$ is singular then it has a zero determinant and the inverse cannot be found in general.

1.10. Moore-Penrose generalized inverse.

$$\mathbf{N}^+ = \mathbf{N}^T (\mathbf{N} \mathbf{N}^T)^{-1}, \quad \mathbf{N} \in \mathbb{R}^{m \times n}, \quad n > m.$$

1.11. Diagonal matrix ($n \times n$ elements). A matrix is diagonal if $N_{ij} = 0$ if $i \neq j$. $\mathbf{N}$ may be written:

$$\mathbf{N} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

The inverse of a square diagonal matrix is $(\mathbf{N}^{-1})_{ii} = (\mathbf{N})_{ii}^{-1}$, $(\mathbf{N}^{-1})_{ij} = 0$ for $i \neq j$:

$$\begin{pmatrix} N_{11} & 0 & \cdots \\ 0 & N_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}^{-1} = \begin{pmatrix} 1/N_{11} & 0 & \cdots \\ 0 & 1/N_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

1.12. Gramian matrix. A Gramian matrix is symmetric and has the form $\mathbf{N}^T \mathbf{N}$:

$$\mathbf{N}^T \mathbf{N} \in \mathbb{R}^{n \times n}, \quad \mathbf{N} \in \mathbb{R}^{m \times n}, \quad \mathbf{N}^T \in \mathbb{R}^{n \times m}.$$

1.13. Euclidean vector inner product (scalar product/dot product) (n elements). Two different vectors:

$$a = \mathbf{v}^a \cdot \mathbf{v}^b = \mathbf{v}^{aT} \mathbf{v}^b = \langle \mathbf{v}^a, \mathbf{v}^b \rangle = \sum_{i=1}^n v_i^a v_i^b \text{ as below:}$$

$$\begin{pmatrix} v_1^a \\ \vdots \\ v_i^a \\ \vdots \\ v_n^a \end{pmatrix} \cdot \begin{pmatrix} v_1^b \\ \vdots \\ v_i^b \\ \vdots \\ v_n^b \end{pmatrix} = \begin{pmatrix} v_1^a & \cdots & v_i^a & \cdots & v_n^a \end{pmatrix} \begin{pmatrix} v_1^b \\ \vdots \\ v_i^b \\ \vdots \\ v_n^b \end{pmatrix} = v_1^a v_1^b + \cdots + v_i^a v_i^b + \cdots + v_n^a v_n^b.$$

The same vector:

$$b = \mathbf{v} \cdot \mathbf{v} = \mathbf{v}^T \mathbf{v} = \langle \mathbf{v}, \mathbf{v} \rangle = \sum_{i=1}^n v_i^2 = \|\mathbf{v}\|^2 \text{ as below:}$$

$$\begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1 & \cdots & v_i & \cdots & v_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix} = v_1 v_1 + \cdots + v_i v_i + \cdots + v_n v_n.$$

1.14. Non-Euclidean vector inner product. Two different vectors (m elements and n elements):

$$a = \mathbf{v}^a \cdot (\mathbf{C} \mathbf{v}^b) = \mathbf{v}^{aT} \mathbf{C} \mathbf{v}^b = \langle \mathbf{v}^a, \mathbf{v}^b \rangle_{\mathbf{C}} = \sum_{i=1}^m v_i^a \sum_{j=1}^n C_{ij} v_j^b \text{ as below:}$$

$$\begin{pmatrix} v_1^a & \cdots & v_i^a & \cdots & v_m^a \end{pmatrix} \begin{pmatrix} C_{11} & \cdots & C_{1j} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ C_{i1} & \cdots & C_{ij} & \cdots & C_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{m1} & \cdots & C_{mj} & \cdots & C_{mn} \end{pmatrix} \begin{pmatrix} v_1^b \\ \vdots \\ v_j^b \\ \vdots \\ v_n^b \end{pmatrix}.$$

The same vector (n elements):

$$b = \mathbf{v} \cdot (\mathbf{C} \mathbf{v}) = \mathbf{v}^T \mathbf{C} \mathbf{v} = \langle \mathbf{v}, \mathbf{v} \rangle_{\mathbf{C}} = \sum_{i=1}^n v_i \sum_{j=1}^n C_{ij} v_j = \|\mathbf{v}\|_{\mathbf{C}}^2 \text{ as below:}$$

$$\begin{pmatrix} v_1 & \cdots & v_i & \cdots & v_n \end{pmatrix} \begin{pmatrix} C_{11} & \cdots & C_{1j} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ C_{i1} & \cdots & C_{ij} & \cdots & C_{in} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{n1} & \cdots & C_{nj} & \cdots & C_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_j \\ \vdots \\ v_n \end{pmatrix}.$$

1.15. Vector outer product ($m \times n$ elements).

$$\mathbf{N} = \mathbf{v}^a \mathbf{v}^b{}^T, \quad N_{ij} = v_i^a v_j^b \text{ as below:}$$

$$\begin{pmatrix} v_1^a \\ \vdots \\ v_i^a \\ \vdots \\ v_m^a \end{pmatrix} \begin{pmatrix} v_1^b & \cdots & v_j^b & \cdots & v_n^b \end{pmatrix} = \begin{pmatrix} v_1^a v_1^b & \cdots & v_1^a v_j^b & \cdots & v_1^a v_n^b \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ v_i^a v_1^b & \cdots & v_i^a v_j^b & \cdots & v_i^a v_n^b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_m^a v_1^b & \cdots & v_m^a v_j^b & \cdots & v_m^a v_n^b \end{pmatrix}.$$

1.16. **Schur (or Hadamard) product.** For matrices ($m \times n$ elements):

$$\mathbf{N} = \mathbf{N}^a \circ \mathbf{N}^b, \quad N_{ij} = N_{ij}^a N_{ij}^b \text{ as below:}$$

$$\begin{pmatrix} N_{11}^a & \cdots & N_{1n}^a \\ \vdots & \ddots & \vdots \\ N_{m1}^a & \cdots & N_{mn}^a \end{pmatrix} \circ \begin{pmatrix} N_{11}^b & \cdots & N_{1n}^b \\ \vdots & \ddots & \vdots \\ N_{m1}^b & \cdots & N_{mn}^b \end{pmatrix} = \begin{pmatrix} N_{11}^a N_{11}^b & \cdots & N_{1n}^a N_{1n}^b \\ \vdots & \ddots & \vdots \\ N_{m1}^a N_{m1}^b & \cdots & N_{mn}^a N_{mn}^b \end{pmatrix}.$$

For vectors (n elements)

$$\mathbf{v} = \mathbf{v}^a \circ \mathbf{v}^b, \quad v_i = v_i^a v_i^b \text{ as below:}$$

$$\begin{pmatrix} v_1^a \\ \vdots \\ v_n^a \end{pmatrix} \circ \begin{pmatrix} v_1^b \\ \vdots \\ v_n^b \end{pmatrix} = \begin{pmatrix} v_1^a v_1^b \\ \vdots \\ v_n^a v_n^b \end{pmatrix}.$$

1.17. **Orthogonal matrix.** If $\mathbf{V}$ is orthogonal then:

$$\mathbf{V}^T \mathbf{V} = \mathbf{I}_n, \quad \mathbf{V} \in \mathbb{R}^{m \times n}, \quad n \leq m.$$

$$\text{If } n = m \text{ then } \mathbf{V}^T = \mathbf{V}^{-1}.$$

1.18. **The trace of a matrix** ($n \times n$ elements). The trace of a square matrix $\mathbf{N}$, $\text{tr}(\mathbf{N})$, is the sum of the diagonal elements:

$$\text{tr}(\mathbf{N}) = \sum_{i=1}^n N_{ii}.$$

1.19. **The Sherman-Morrison-Woodbury formula.**

$$(\mathbf{A} + \mathbf{C}\mathbf{D}^T)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{I} + \mathbf{D}^T\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{D}^T\mathbf{A}^{-1}.$$

Replacing $\mathbf{C} \rightarrow \mathbf{C}\mathbf{B}$ and then setting $\mathbf{C} = \mathbf{D} = \mathbf{H}$ and $\mathbf{A} = \mathbf{R}$, the following useful formula results:

$$(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})\mathbf{B}\mathbf{H}^T = \mathbf{H}^T\mathbf{R}^{-1}(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T).$$

2. FUNCTIONS

2.1. **Scalar valued function of a vector** (n elements) **and its derivative.**

$$f(\mathbf{v}), \quad f, \quad \nabla_{\mathbf{v}} f(\mathbf{v}) = \left(\frac{\partial f}{\partial \mathbf{v}} \right)^T = \begin{pmatrix} \partial f / \partial v_1 \\ \partial f / \partial v_2 \\ \vdots \\ \partial f / \partial v_n \end{pmatrix}.$$

2.2. **Generalised chain rule.**

$$\text{Consider } f(\mathbf{v}^b), \text{ where } \nabla_{\mathbf{v}^b} f(\mathbf{v}^b) \text{ is known, } f \in \mathbb{R}, \quad \mathbf{v}^b, \nabla_{\mathbf{v}^b} f(\mathbf{v}^b) \in \mathbb{R}^m.$$

$$\text{If } \mathbf{v}^b = \mathbf{N}\mathbf{v}^a, \text{ then } \nabla_{\mathbf{v}^a} f(\mathbf{v}^a) = \mathbf{N}^T \nabla_{\mathbf{v}^b} f(\mathbf{v}^b), \quad \mathbf{v}^a, \nabla_{\mathbf{v}^a} f(\mathbf{v}^a) \in \mathbb{R}^n, \quad \mathbf{N} \in \mathbb{R}^{m \times n}.$$

2.3. Generalised Taylor series for f . Let $f(\mathbf{v})$ be a linear or non-linear function. The Taylor series of $f(\mathbf{v})$ about $\mathbf{v}$ is:

$$f(\mathbf{v} + \delta\mathbf{v}) = f(\mathbf{v}) + \frac{\partial f}{\partial \mathbf{v}} \delta\mathbf{v} + \frac{1}{2} \delta\mathbf{v}^T \frac{\partial^2 f}{\partial \mathbf{v}^2} \delta\mathbf{v} + \text{higher order terms},$$

$$f \in \mathbb{R}, \quad \mathbf{v}, \frac{\partial f}{\partial \mathbf{v}} \in \mathbb{R}^n, \quad \frac{\partial^2 f}{\partial \mathbf{v}^2} \in \mathbb{R}^{n \times n} \text{ is the Hessian matrix, } \left(\frac{\partial^2 f}{\partial \mathbf{v}^2} \right)_{ij} = \frac{\partial^2 f}{\partial v_i \partial v_j}.$$

2.4. Vector valued function of a vector.

$$\mathbf{f}(\mathbf{v}), \quad \mathbf{f} \in \mathbb{R}^m, \quad \mathbf{v} \in \mathbb{R}^n.$$

2.5. Generalised Taylor series for $\mathbf{f}$. Let $\mathbf{f}(\mathbf{v})$ be a linear or non-linear function. The Taylor series of $\mathbf{f}(\mathbf{v})$ about $\mathbf{v}$ is:

$$\mathbf{f}(\mathbf{v} + \delta\mathbf{v}) = \mathbf{f}(\mathbf{v}) + \mathbf{F} \delta\mathbf{v} + \text{higher order terms},$$

$$\mathbf{F} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{v}} \right|_{\mathbf{v}}, \quad F_{ij} = \left. \frac{\partial f_i}{\partial v_j} \right|_{\mathbf{v}}, \quad \mathbf{f} \in \mathbb{R}^m, \quad \mathbf{v} \in \mathbb{R}^n, \quad \mathbf{F} \in \mathbb{R}^{m \times n}.$$

$\mathbf{F}$ is the Jacobian of $\mathbf{f}(\mathbf{v})$ about $\mathbf{v}$ and $\partial f_i / \partial v_j$ are called Fréchet derivatives.

3. MATRIX DECOMPOSITIONS

3.1. Eigenvectors and eigenvalues. The k th eigenvector ($\mathbf{v}_k$) and eigenvalue (λ_k) of matrix $\mathbf{N}$ satisfies:

$$\mathbf{N} \mathbf{v}_k = \lambda_k \mathbf{v}_k, \quad \mathbf{N} \in \mathbb{R}^{n \times n}, \quad \mathbf{v}_k \in \mathbb{R}^n, \quad \lambda_k \in \mathbb{R}, \quad 1 \leq k \leq n.$$

$$\text{Let } \mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} \begin{pmatrix} \mathbf{v}_2 \\ \mathbf{v}_1 \end{pmatrix} \cdots \begin{pmatrix} \mathbf{v}_n \\ \mathbf{v}_1 \end{pmatrix}, \quad \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

$$\mathbf{N} \mathbf{V} = \mathbf{V} \Lambda, \quad \mathbf{N}, \mathbf{V}, \Lambda \in \mathbb{R}^{n \times n}.$$

If $\mathbf{N}$ is Hermitian (if a real matrix then this is equivalent to $\mathbf{N}$ being symmetric) then $\mathbf{V}$ (the matrix of eigenvectors) is orthogonal (see below), and Λ (the matrix of eigenvalues) is real.

For a general 2×2 matrix:

$$\mathbf{N} = \begin{pmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \alpha_1 \gamma_1 & \alpha_2 \gamma_2 \\ \alpha_1 & \alpha_2 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \frac{N_{11} + N_{22} - \beta}{2} & 0 \\ 0 & \frac{N_{11} + N_{22} + \beta}{2} \end{pmatrix},$$

$$\beta = \sqrt{N_{11}^2 - 2N_{11}N_{22} + 4N_{12}N_{21} + N_{22}^2},$$

$$\gamma_1 = \frac{N_{11} - N_{22} - \beta}{2N_{21}}, \quad \gamma_2 = \frac{N_{11} - N_{22} + \beta}{2N_{21}}, \quad \alpha_1 = \frac{1}{\sqrt{\gamma_1^2 + 1}}, \quad \alpha_2 = \frac{1}{\sqrt{\gamma_2^2 + 1}}.$$

3.2. Singular vectors and singular values.

$$\mathbf{N} \mathbf{V} = \mathbf{U} \Lambda, \quad \mathbf{N}^T \mathbf{U} = \mathbf{V} \Lambda, \quad \mathbf{U}^T \mathbf{U} = \mathbf{I}_p, \quad \mathbf{V}^T \mathbf{V} = \mathbf{I}_p.$$

$$\mathbf{N} \in \mathbb{R}^{m \times n}, \quad \mathbf{V} \in \mathbb{R}^{n \times p}, \quad \mathbf{U} \in \mathbb{R}^{m \times p}, \quad \Lambda \in \mathbb{R}^{p \times p}, \quad p = \text{rank of } \mathbf{N}.$$

$\mathbf{V}$ is the matrix of right singular vectors of $\mathbf{N}$, $\mathbf{U}$ is the matrix of left singular vectors of $\mathbf{N}$, and Λ is the matrix of singular values of $\mathbf{N}$. The following eigenvalue equations exist for $\mathbf{V}$ and $\mathbf{U}$:

$$\mathbf{N}^T \mathbf{N} \mathbf{V} = \mathbf{V} \Lambda, \quad \mathbf{N} \mathbf{N}^T \mathbf{U} = \mathbf{U} \Lambda.$$

3.3. The rank of a matrix. The rank of $\mathbf{N}$ is the number of independent rows or columns of $\mathbf{N}$ (consider, e.g. the i th column of $\mathbf{N}$ as vector $\mathbf{n}_i$). A column (or row) is dependent if it can be written as a linear combination of the other columns (or rows). The rank of a matrix is also the number of non-zero singular values of $\mathbf{N}$. The rank of a square matrix is also the number of non-zero eigenvalues.

4. MEAN, (CO)VARIANCE, CORRELATION AND GAUSSIAN STATISTICS

4.1. The variance, standard deviation and mean of a scalar. Consider a population of N scalars, s^l , $1 \leq l \leq N$. The following are for the variance, $\text{var}(s)$, standard deviation, σ_s , and mean, $\langle s \rangle$ (common notations are given)¹:

$$\text{var}(s) = \langle (s - \langle s \rangle)^2 \rangle = \overline{(s - \bar{s})^2} = \mathcal{E}((s - \mathcal{E}(s))^2) \approx \frac{1}{\tilde{N}} \sum_{l=1}^N (s^l - \langle s \rangle)^2,$$

$$\sigma_s = \sqrt{\text{var}(s)}, \quad \langle s \rangle = \bar{s} = \mathcal{E}(s) \approx \frac{1}{\tilde{N}} \sum_{l=1}^N s^l.$$

4.2. The covariance between two scalars. Consider two populations, each of N scalars, s^l , t^l , $1 \leq l \leq N$. The following is for the covariance, $\text{cov}(s, t)$ (common notations are given)²:

$$\text{cov}(s, t) = \langle (s - \langle s \rangle)(t - \langle t \rangle) \rangle = \overline{(s - \bar{s})(t - \bar{t})} = \mathcal{E}((s - \mathcal{E}(s))(t - \mathcal{E}(t))) \approx \frac{1}{\tilde{N}} \sum_{l=1}^N (s^l - \langle s \rangle)(t^l - \langle t \rangle).$$

The covariance between two scalars can be negative, zero or positive.

4.3. The correlation between two scalars.

$$\text{cor}(s, t) = \frac{\text{cov}(s, t)}{\sigma_s \sigma_t}, \quad -1 \leq \text{cor}(s, t) \leq 1, \quad \text{cor}(s, s) = \text{cor}(t, t) = 1.$$

4.4. The covariance matrix between two vectors (one with m elements, another with n elements). Consider two populations, each of N vectors, $\mathbf{u}^l$, $\mathbf{v}^l$, $1 \leq l \leq N$. The following is for the covariance matrix, $\text{cov}(\mathbf{u}, \mathbf{v})$, which uses the outer product (common notations are given):

$$\begin{aligned} \text{cov}(\mathbf{u}, \mathbf{v}) &= \langle (\mathbf{u} - \langle \mathbf{u} \rangle)(\mathbf{v} - \langle \mathbf{v} \rangle)^T \rangle = \overline{(\mathbf{u} - \bar{\mathbf{u}})(\mathbf{v} - \bar{\mathbf{v}})^T} = \mathcal{E}((\mathbf{u} - \mathcal{E}(\mathbf{u}))(\mathbf{v} - \mathcal{E}(\mathbf{v}))^T), \\ &\approx \frac{1}{\tilde{N}} \sum_{l=1}^N (\mathbf{u}^l - \langle \mathbf{u} \rangle)(\mathbf{v}^l - \langle \mathbf{v} \rangle)^T, \\ (\text{cov}(\mathbf{u}, \mathbf{v}))_{ij} &\approx \frac{1}{\tilde{N}} \sum_{l=1}^N (u_i^l - \langle u_i \rangle)(v_j^l - \langle v_j \rangle), \\ \text{cov}(\mathbf{u}, \mathbf{v}) &= \frac{1}{\tilde{N}} \sum_{l=1}^N \begin{pmatrix} (u_1^l - \langle u_1 \rangle)(v_1^l - \langle v_1 \rangle) & \cdots & (u_1^l - \langle u_1 \rangle)(v_j^l - \langle v_j \rangle) & \cdots & (u_1^l - \langle u_1 \rangle)(v_n^l - \langle v_n \rangle) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ (u_i^l - \langle u_i \rangle)(v_1^l - \langle v_1 \rangle) & \cdots & (u_i^l - \langle u_i \rangle)(v_j^l - \langle v_j \rangle) & \cdots & (u_i^l - \langle u_i \rangle)(v_n^l - \langle v_n \rangle) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ (u_m^l - \langle u_m \rangle)(v_1^l - \langle v_1 \rangle) & \cdots & (u_m^l - \langle u_m \rangle)(v_j^l - \langle v_j \rangle) & \cdots & (u_m^l - \langle u_m \rangle)(v_n^l - \langle v_n \rangle) \end{pmatrix}. \end{aligned}$$

If $\mathbf{u} = \mathbf{v}$, then $\text{cov}(\mathbf{v}, \mathbf{v})$ is the auto-covariance matrix of $\mathbf{v}$. Diagonal elements are variances of each element of $\mathbf{v}$, i.e. $(\text{cov}(\mathbf{v}, \mathbf{v}))_{ii} = \text{var}(v_i)$:

$$\text{cov}(\mathbf{v}, \mathbf{v}) = \frac{1}{\tilde{N}} \sum_{l=1}^N \begin{pmatrix} (v_1^l - \langle v_1 \rangle)^2 & \cdots & (v_1^l - \langle v_1 \rangle)(v_i^l - \langle v_i \rangle) & \cdots & (v_1^l - \langle v_1 \rangle)(v_n^l - \langle v_n \rangle) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ (v_i^l - \langle v_i \rangle)(v_1^l - \langle v_1 \rangle) & \cdots & (v_i^l - \langle v_i \rangle)^2 & \cdots & (v_i^l - \langle v_i \rangle)(v_n^l - \langle v_n \rangle) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ (v_n^l - \langle v_n \rangle)(v_1^l - \langle v_1 \rangle) & \cdots & (v_n^l - \langle v_n \rangle)(v_i^l - \langle v_i \rangle) & \cdots & (v_n^l - \langle v_n \rangle)^2 \end{pmatrix}.$$

¹Sample variance and sample mean are those (approximate expressions) given in terms of a finite population (a sample). In the formula for the sample variance, $\tilde{N} = N$ if $\langle s \rangle$ is taken to be the *exact* mean, but $\tilde{N} = N - 1$ if $\langle s \rangle$ is taken to be the *sample* mean.

²Sample covariance and sample mean are those (approximate expressions) given in terms of a finite population (a sample). In the formula for the sample covariance, $\tilde{N} = N$ if $\langle s \rangle$ and $\langle t \rangle$ are taken to be the *exact* means, but $\tilde{N} = N - 1$ if $\langle s \rangle$ and $\langle t \rangle$ are taken to be the *sample* means.

Auto-covariance matrices are symmetric.

4.5. The correlation matrix between two vectors.

$\text{cor}(\mathbf{u}, \mathbf{v}) = \Sigma_{\mathbf{u}}^{-1} \text{cov}(\mathbf{u}, \mathbf{v}) \Sigma_{\mathbf{v}}^{-1}$, $\Sigma_{\mathbf{u}} = \text{diag}(\sigma_{u_1}, \sigma_{u_2}, \dots, \sigma_{u_m})$, $\Sigma_{\mathbf{v}} = \text{diag}(\sigma_{v_1}, \sigma_{v_2}, \dots, \sigma_{v_n})$ as below:

$$\text{cor}(\mathbf{u}, \mathbf{v}) = \begin{pmatrix} \sigma_{u_1}^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_{u_2}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{u_m}^{-1} \end{pmatrix} \begin{pmatrix} (\text{cov}(\mathbf{u}, \mathbf{v}))_{11} & (\text{cov}(\mathbf{u}, \mathbf{v}))_{12} & \cdots & (\text{cov}(\mathbf{u}, \mathbf{v}))_{1n} \\ (\text{cov}(\mathbf{u}, \mathbf{v}))_{21} & (\text{cov}(\mathbf{u}, \mathbf{v}))_{22} & \cdots & (\text{cov}(\mathbf{u}, \mathbf{v}))_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ (\text{cov}(\mathbf{u}, \mathbf{v}))_{m1} & (\text{cov}(\mathbf{u}, \mathbf{v}))_{m2} & \cdots & (\text{cov}(\mathbf{u}, \mathbf{v}))_{mn} \end{pmatrix} \times \begin{pmatrix} \sigma_{v_1}^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_{v_2}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{v_n}^{-1} \end{pmatrix},$$

$$\text{i.e. } (\text{cor}(\mathbf{u}, \mathbf{v}))_{ij} = \frac{(\text{cov}(\mathbf{u}, \mathbf{v}))_{ij}}{\sigma_{u_i} \sigma_{v_j}}.$$

If $\mathbf{u} = \mathbf{v}$ then $\text{cor}(\mathbf{v}, \mathbf{v})$ is the auto-correlation matrix of $\mathbf{v}$:

$$\text{cor}(\mathbf{v}, \mathbf{v}) = \Sigma_{\mathbf{v}}^{-1} \text{cov}(\mathbf{v}, \mathbf{v}) \Sigma_{\mathbf{v}}^{-1} \text{ as below:}$$

$$\text{cor}(\mathbf{v}, \mathbf{v}) = \begin{pmatrix} \sigma_{v_1}^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_{v_2}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{v_n}^{-1} \end{pmatrix} \begin{pmatrix} (\text{cov}(\mathbf{v}, \mathbf{v}))_{11} & (\text{cov}(\mathbf{v}, \mathbf{v}))_{12} & \cdots & (\text{cov}(\mathbf{v}, \mathbf{v}))_{1n} \\ (\text{cov}(\mathbf{v}, \mathbf{v}))_{21} & (\text{cov}(\mathbf{v}, \mathbf{v}))_{22} & \cdots & (\text{cov}(\mathbf{v}, \mathbf{v}))_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ (\text{cov}(\mathbf{v}, \mathbf{v}))_{n1} & (\text{cov}(\mathbf{v}, \mathbf{v}))_{n2} & \cdots & (\text{cov}(\mathbf{v}, \mathbf{v}))_{nn} \end{pmatrix} \times, \begin{pmatrix} \sigma_{v_1}^{-1} & 0 & \cdots & 0 \\ 0 & \sigma_{v_2}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{v_n}^{-1} \end{pmatrix},$$

$$= \begin{pmatrix} 1 & (\text{cor}(\mathbf{v}, \mathbf{v}))_{12} & \cdots & (\text{cor}(\mathbf{v}, \mathbf{v}))_{1n} \\ (\text{cor}(\mathbf{v}, \mathbf{v}))_{21} & 1 & \cdots & (\text{cor}(\mathbf{v}, \mathbf{v}))_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ (\text{cor}(\mathbf{v}, \mathbf{v}))_{n1} & (\text{cor}(\mathbf{v}, \mathbf{v}))_{n2} & \cdots & 1 \end{pmatrix},$$

$$\text{i.e. } (\text{cor}(\mathbf{v}, \mathbf{v}))_{ij} = \frac{(\text{cov}(\mathbf{v}, \mathbf{v}))_{ij}}{\sigma_{v_i} \sigma_{v_j}}.$$

Auto-correlation matrices are symmetric.

4.6. Gaussian/normal probability density function. For n -variables:

$$p(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\mathbf{P})}} \exp \left[-\frac{1}{2} (\mathbf{x} - \langle \mathbf{x} \rangle)^T \mathbf{P}^{-1} (\mathbf{x} - \langle \mathbf{x} \rangle) \right], \quad \mathbf{P} = \text{cov}(\mathbf{x}, \mathbf{x}).$$

For $n = 1$:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(x - \langle x \rangle)^2}{2\sigma^2} \right].$$

For $n = 2$:

$$p(\mathbf{x}) = p(x_1, x_2) = \frac{1}{\sqrt{4\pi^2(\sigma_1^2\sigma_2^2 - P_{12}^2)}} \exp \left[-\frac{1}{2} \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & P_{12} \\ P_{12} & \sigma_2^2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right].$$

5. FOURIER ANALYSIS

5.1. The Fourier transform. The real-to-spectral space transform in 1-D (1-D Fourier transform):

$$\bar{f}(k) = \frac{1}{\sqrt{2\pi}} \int f(x) \exp(-ikx) dx, \quad i = \sqrt{-1}.$$

The spectral-to-real transform in 1-D (1-D inverse Fourier transform):

$$f(x) = \frac{1}{\sqrt{2\pi}} \int \bar{f}(k) \exp(ikx) dk.$$

The real-to-spectral space transform in d dimensions:

$$\bar{f}(\mathbf{k}) = \frac{1}{(2\pi)^{d/2}} \int \int \int f(\mathbf{x}) \exp(-i\mathbf{k} \cdot \mathbf{x}) d\mathbf{x}.$$

The spectral-to-real transform in d dimensions:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}} \int \int \int \bar{f}(\mathbf{k}) \exp(i\mathbf{k} \cdot \mathbf{x}) d\mathbf{k}.$$

The Fourier transforms rely on the orthogonality relationships:

$$\begin{aligned} \int \int \int \exp(i\mathbf{k} \cdot \mathbf{x}) \exp(-i\mathbf{k}' \cdot \mathbf{x}) d\mathbf{x} &= (2\pi)^d \delta(\mathbf{k} - \mathbf{k}'), \\ \int \int \int \exp(i\mathbf{k} \cdot \mathbf{x}) \exp(-i\mathbf{k} \cdot \mathbf{x}') d\mathbf{k} &= (2\pi)^d \delta(\mathbf{x} - \mathbf{x}'), \end{aligned}$$

and satisfies the convolution theorem:

$$\int g(x - x') f(x') dx' \quad \text{has Fourier transform} \quad 2\pi \bar{g}(k) \bar{f}(k).$$

5.2. Fourier series. Fourier series are the discrete versions of the Fourier transforms (real and spectral spaces comprising N discrete points). In 1-D:

$$\begin{aligned} \bar{f}(k_i) &= \frac{1}{\sqrt{N}} \sum_{j=1}^N f(x_j) \exp(-ik_i x_j), \quad f(x_j) = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} \bar{f}(k_i) \exp(ik_i x_j), \\ \sum_{j=1}^N \exp(ik_i x_j) \exp(ik_{i'} x_j) &= N \delta_{ii'}, \quad \sum_{i=0}^{N-1} \exp(ik_i x_j) \exp(ik_{i'} x_j) = N \delta_{jj'}. \end{aligned}$$

Representing $f(x_j)$ as the vector $\mathbf{f}$ and $\bar{f}(k_i)$ as the vector $\bar{\mathbf{f}}$ allows the discrete Fourier series, its inverse, and the orthogonality relations to be written compactly via an orthogonal matrix transform:

$$\bar{\mathbf{f}} = \mathbf{F}\mathbf{f}, \quad \mathbf{f} = \mathbf{F}^\dagger \bar{\mathbf{f}}, \quad \mathbf{F}^\dagger \mathbf{F} = \mathbf{I}_N, \quad \mathbf{F}\mathbf{F}^\dagger = \mathbf{I}_N, \quad \text{where matrix elements } F_{ij} = \frac{1}{\sqrt{N}} \exp(-ik_i x_j).$$

6. VARIATIONAL CALCULUS

6.1. Lagrange multipliers. Problem: find the stationary point of $f(x_1, x_2, \dots, x_N)$ subject to the constraint $g_m(x_1, x_2, \dots, x_N)$, $1 \leq m \leq M$. This problem has N degrees of freedom and M constraints. The constrained variational problem can be written as (f and g_m are implied functions of $x_1, x_2, \dots, x_N$):

$$\frac{\partial}{\partial x_n} \left(f + \sum_{m=1}^M g_m \lambda_m \right) = 0, \quad 1 \leq n \leq N,$$

where λ_m is the Lagrange multiplier associated with the m th constraint. This can be written in the following matrix form:

$$\nabla_{\mathbf{x}} f + \mathbf{G}^T \boldsymbol{\lambda} = 0, \quad \mathbf{x} \in \mathbb{R}^N, \quad \boldsymbol{\lambda} \in \mathbb{R}^M, \quad \mathbf{G} \in \mathbb{R}^{M \times N},$$

where $\mathbf{x} = (x_1, x_2, \dots, x_N)^T$ and $G_{mn} = \partial g_m / \partial x_n$.
