Summary of Mathematical Tools useful for Data Assimilation Theory

Vector representation of information

$$\mathbf{v} \in \mathbb{R}^n.$$

$$\mathbf{v} = \begin{array}{c} \\ i \\ \\ \end{array}\begin{pmatrix} \\ \mathbf{v}_i \\ \\ \end{pmatrix}$$

Matrix operator

$$\mathbf{v}_\mathrm{B} = \mathbf{N}\mathbf{v}_\mathrm{A},$$

$$\mathbf{v}_\mathrm{B} \in \mathbb{R}^m, \; \mathbf{v}_\mathrm{A} \in \mathbb{R}^n, \; \mathbf{N} \in \mathbb{R}^{m \times n},$$

$$(\mathbf{v}_\mathrm{B})_i = \sum_{j=1}^{n} \mathbf{N}_{ij}\,(\mathbf{v}_\mathrm{A})_j, \; 1 \leqslant i \leqslant m.$$

$$\mathbf{N} = \begin{array}{cc} & \begin{array}{ccc} \quad & j & \quad\quad\quad \end{array} \\ i & \begin{pmatrix} & & \\ & \mathbf{N}_{ij} & \\ & & \end{pmatrix} \end{array}$$

Identity (or unit) matrix

$$\mathbf{I}_p = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}, \quad (\mathbf{I}_p)_{ij} = \delta_{ij}, \quad \mathbf{I}_p \in \mathbb{R}^{p \times p}.$$

Matrix addition

$$\mathbf{N} = \mathbf{N}_\mathrm{A} + \mathbf{N}_\mathrm{B}, \quad \mathbf{N}_{ij} = (\mathbf{N}_\mathrm{A})_{ij} + (\mathbf{N}_\mathrm{B})_{ij},$$

$$\mathbf{N}, \; \mathbf{N}_\mathrm{A}, \; \mathbf{N}_\mathrm{B} \in \mathbb{R}^{m \times n}.$$

Matrix multiplication

$$\mathbf{N} = \mathbf{N}_\mathrm{A}\mathbf{N}_\mathrm{B}, \quad \mathbf{N}_{ij} = \sum_{k=1}^{p} (\mathbf{N}_\mathrm{A})_{ik}\,(\mathbf{N}_\mathrm{B})_{kj},$$

$$\mathbf{N} \in \mathbb{R}^{m \times n}, \; \mathbf{N}_\mathrm{A} \in \mathbb{R}^{m \times p}, \; \mathbf{N}_\mathrm{B} \in \mathbb{R}^{p \times n}.$$

$\mathbf{N}_\mathrm{A}$ and $\mathbf{N}_\mathrm{B}$ are not commutative, i.e. in general

$$\mathbf{N}_\mathrm{A}\mathbf{N}_\mathrm{B} \neq \mathbf{N}_\mathrm{B}\mathbf{N}_\mathrm{A}.$$

Pre multiplication by the identity matrix

$$\mathbf{I}_p\mathbf{N}_\mathrm{B} = \mathbf{N}_\mathrm{B}.$$

Post multiplication by the identity matrix

$$\mathbf{N}_\mathrm{A}\mathbf{I}_p = \mathbf{N}_\mathrm{A}.$$

Multiplication by a scalar

$$(\alpha\mathbf{N})_{ij}, \; \alpha\,\mathbf{N}_{ij}.$$

Matrix inversion

Let $\mathbf{N}$ be a square ($m = n$) non-singular matrix.

$$\text{If } \mathbf{v}_\mathrm{B} = \mathbf{N}\mathbf{v}_\mathrm{A}, \text{ then } \mathbf{v}_\mathrm{A} = \mathbf{N}^{-1}\mathbf{v}_\mathrm{B},$$

$$\mathbf{v}_\mathrm{A}, \; \mathbf{v}_\mathrm{B} \in \mathbb{R}^n, \; \mathbf{N}, \mathbf{N}^{-1} \in \mathbb{R}^{n \times n}.$$

$\mathbf{N}$ is singular if it has zero determinant, and the inverse cannot be found in general. In general

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

Computing the inverse of a matrix is a non-trivial exercise in general. For a

2 × 2 matrix

$$\mathbf{N} = \begin{pmatrix} \mathbf{N}_{11} & \mathbf{N}_{12} \\ \mathbf{N}_{21} & \mathbf{N}_{22} \end{pmatrix}, \qquad \mathbf{N}^{-1} = \frac{1}{\det(\mathbf{N})}\begin{pmatrix} \mathbf{N}_{22} & -\mathbf{N}_{12} \\ -\mathbf{N}_{21} & \mathbf{N}_{11} \end{pmatrix},$$

$$\det(\mathbf{N}) = \mathbf{N}_{11}\mathbf{N}_{22} - \mathbf{N}_{12}\mathbf{N}_{21}.$$

Diagonal matrix

A diagonal matrix has

$$\mathbf{N}_{ij} = 0 \text{ if } i \neq j, \quad \mathbf{N} \in \mathbb{R}^{m \times n}.$$

If **N** is square

$$\mathbf{N} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots) = \begin{pmatrix} \lambda_1 & 0 & \ldots \\ 0 & \lambda_2 & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}.$$

The inverse of a square diagonal matrix is

$$(\mathbf{N}^{-1})_{ii} = (\mathbf{N}_{ii})^{-1}, \quad (\mathbf{N}^{-1})_{ij} = 0 \text{ for } i \neq j,$$

$$\begin{pmatrix} \mathbf{N}_{11} & 0 & 0 & \ldots \\ 0 & \mathbf{N}_{22} & 0 & \ldots \\ 0 & 0 & \mathbf{N}_{33} & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{pmatrix}^{-1} = \begin{pmatrix} 1/\mathbf{N}_{11} & 0 & 0 & \ldots \\ 0 & 1/\mathbf{N}_{22} & 0 & \ldots \\ 0 & 0 & 1/\mathbf{N}_{33} & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{pmatrix}.$$

Matrix transpose

The matrix transpose makes rows into columns and columns into rows

$$\mathbf{N}_\mathrm{B} = \mathbf{N}_\mathrm{A}^\mathrm{T}, \quad (\mathbf{N}_\mathrm{B})_{ij} = (\mathbf{N}_\mathrm{A})_{ji},$$

$$\mathbf{N}_\mathrm{B} \in \mathbb{R}^{m \times n}, \ \mathbf{N}_\mathrm{A} \in \mathbb{R}^{n \times m},$$

$$\mathbf{N}_\mathrm{A} = \begin{pmatrix} (\mathbf{N}_\mathrm{A})_{11} & (\mathbf{N}_\mathrm{A})_{12} & (\mathbf{N}_\mathrm{A})_{13} \\ (\mathbf{N}_\mathrm{A})_{21} & (\mathbf{N}_\mathrm{A})_{22} & (\mathbf{N}_\mathrm{A})_{23} \end{pmatrix}, \qquad \mathbf{N}_\mathrm{B} = \begin{pmatrix} (\mathbf{N}_\mathrm{A})_{11} & (\mathbf{N}_\mathrm{A})_{21} \\ (\mathbf{N}_\mathrm{A})_{12} & (\mathbf{N}_\mathrm{A})_{22} \\ (\mathbf{N}_\mathrm{A})_{13} & (\mathbf{N}_\mathrm{A})_{23} \end{pmatrix}.$$

Transpose of a product of matrices

$$(\mathbf{N}_\mathrm{A}\mathbf{N}_\mathrm{B})^\mathrm{T} = \mathbf{N}_\mathrm{B}^\mathrm{T}\mathbf{N}_\mathrm{A}^\mathrm{T}.$$

Symmetric matrix

A matrix is symmetric if

$$\mathbf{N} = \mathbf{N}^\mathrm{T}, \ \mathbf{N}_{ij} = \mathbf{N}_{ji}, \ \mathbf{N} \in \mathbb{R}^{n \times n}.$$

Only square matrices can be symmetric.

Gramian matrices

Gramian matrices are of the form $\mathbf{N}^\mathrm{T}\mathbf{N}$

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

Gramian matrices are symmetric.

Vector inner product (or scalar product, or dot product) (Euclidean)

$$a = \mathbf{v}_\mathrm{A} \cdot \mathbf{v}_\mathrm{B} = \mathbf{v}_\mathrm{A}^\mathrm{T}\mathbf{v}_\mathrm{B} = \langle \mathbf{v}_\mathrm{A}, \mathbf{v}_\mathrm{B} \rangle = \sum_{i=1}^{n} (\mathbf{v}_\mathrm{A})_i (\mathbf{v}_\mathrm{B})_i,$$

$$\mathbf{v}_\mathrm{A}, \mathbf{v}_\mathrm{B} \in \mathbb{R}^n, \ a \in \mathbb{R}^1.$$

$$b = \mathbf{v}^\mathrm{T}\mathbf{v} = \langle \mathbf{v}, \mathbf{v} \rangle = \sum_{i=1}^{n} (\mathbf{v}_i)^2 = \|\mathbf{v}\|^2,$$

$$\mathbf{v} \in \mathbb{R}^n, \ b \in \mathbb{R}^1.$$

Vector inner product (non-Euclidean)

$$a = \mathbf{v}_A \cdot (\mathbf{C}\mathbf{v}_B) = \mathbf{v}_A^T\mathbf{C}\mathbf{v}_B = \langle \mathbf{v}_A, \mathbf{v}_B \rangle_{\mathbf{C}} = \sum_{i=1}^{n} (\mathbf{v}_A)_i \sum_{j=1}^{m} \mathbf{C}_{ij} (\mathbf{v}_B)_j ,$$

$$\mathbf{v}_A \in \mathbb{R}^n, \ \mathbf{v}_B \in \mathbb{R}^m, \ \mathbf{C} \in \mathbb{R}^{n \times m}, \ a \in \mathbb{R}^1,$$

$$b = \mathbf{v}^T\mathbf{C}\mathbf{v} = \langle \mathbf{v}, \mathbf{v} \rangle_{\mathbf{C}} = \sum_{i=1}^{n} \mathbf{v}_i \sum_{j=1}^{m} \mathbf{C}_{ij}\mathbf{v}_j = \| \mathbf{v} \|_{\mathbf{C}}^2,$$

$$\mathbf{v} \in \mathbb{R}^n, \ \mathbf{C} \in \mathbb{R}^{n \times n}, \ b \in \mathbb{R}^1$$

Vector outer product

$$\mathbf{N} = \mathbf{v}_A\mathbf{v}_B^T, \quad \mathbf{N}_{ij} = (\mathbf{v}_A)_i (\mathbf{v}_B)_j,$$

$$\mathbf{N} \in \mathbb{R}^{m \times n}, \ \mathbf{v}_A \in \mathbb{R}^m, \ \mathbf{v}_B \in \mathbb{R}^n.$$

Schur (or Hadamard) product

$$\mathbf{N} = \mathbf{N}_A \circ \mathbf{N}_B, \quad \mathbf{N}_{ij} = (\mathbf{N}_A)_{ij} (\mathbf{N}_B)_{ij} ,$$

$$\mathbf{N}, \ \mathbf{N}_A, \ \mathbf{N}_B \in \mathbb{R}^{m \times n}.$$

Scalar valued function of a vector

$$f(\mathbf{v}), \quad f \in \mathbb{R}^1, \ \mathbf{v} \in \mathbb{R}^n.$$

The vector derivative

$$\nabla_{\mathbf{v}} f(\mathbf{v}) = \left(\frac{\partial f}{\partial \mathbf{v}}\right)^T = \begin{pmatrix} \partial f / \partial \mathbf{v}_1 \\ \partial f / \partial \mathbf{v}_2 \\ \cdots \\ \partial f / \partial \mathbf{v}_n \end{pmatrix},$$

$$f \in \mathbb{R}^1, \ \mathbf{v}, \ \nabla_{\mathbf{v}} f \in \mathbb{R}^n.$$

Generalized chain rule

$$f(\mathbf{v}_B), \qquad \mathbf{v}_B = \mathbf{N}\mathbf{v}_A, \qquad \nabla_{\mathbf{v}_A} f = \mathbf{N}^T \nabla_{\mathbf{v}_B} f,$$

$$f \in \mathbb{R}^1, \mathbf{v}_B \in \mathbb{R}^m, \ \mathbf{v}_A \in \mathbb{R}^n, \ \mathbf{N} \in \mathbb{R}^{m \times n},$$

$$\nabla_{\mathbf{v}_A} f \in \mathbb{R}^n, \ \nabla_{\mathbf{v}_B} f \in \mathbb{R}^m, \ \mathbf{N}^T \in \mathbb{R}^{n \times m}.$$

Vector valued function of a vector

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

The Jacobian of a function

Let $\mathbf{f}(\mathbf{v})$ be a non-linear function of $\mathbf{v}$. The Taylor expansion of $\mathbf{f}(\mathbf{v})$ about $\mathbf{v}$ is

$$\mathbf{f}(\mathbf{v} + \delta\mathbf{v}) = \mathbf{f}(\mathbf{v}) + \mathbf{F}\delta\mathbf{v} + \text{h.o.t.},$$

h.o.t. stands for higher order terms. $\mathbf{F}$ is called the Jacobian of $\mathbf{f}(\mathbf{v})$ about $\mathbf{v}$

$$\mathbf{F} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{v}}\right|_{\mathbf{v}}, \quad \mathbf{F}_{ij} = \left.\frac{\partial \mathbf{f}_i}{\partial \mathbf{v}_j}\right|_{\mathbf{v}},$$

$$\mathbf{f} \in \mathbb{R}^m, \ \mathbf{v}, \delta\mathbf{v} \in \mathbb{R}^n, \ \mathbf{F} \in \mathbb{R}^{m \times n}.$$

$\partial \mathbf{f}_i / \partial \mathbf{v}_j$ are called Fréchet derivatives.

Eigenvectors and eigenvalues

The kth eigenvector ($\mathbf{v}_k$) and eigenvalue (λ_k) of matrix $\mathbf{N}$

$$\mathbf{N}\mathbf{v}_k = \lambda_k\mathbf{v}_k,$$

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

$$\text{Let } \mathbf{V} = (\mathbf{v}_1\, \mathbf{v}_2 \ldots \mathbf{v}_n) = \begin{pmatrix} \\ \mathbf{v}_1 \\ \\ \end{pmatrix}\begin{pmatrix} \\ \mathbf{v}_2 \\ \\ \end{pmatrix} \ldots \begin{pmatrix} \\ \mathbf{v}_n \\ \\ \end{pmatrix},$$

$$\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots \lambda_n),$$

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

If $\mathbf{N}$ is real-valued and symmetric, then $\mathbf{V}$ is an orthonormal matrix. For the $\mathbb{R}^{2 \times 2}$ matrix below the eigenvalues and eigenvectors are

$$\mathbf{N} = \begin{pmatrix} \mathbf{N}_{11} & \mathbf{N}_{12} \\ \mathbf{N}_{21} & \mathbf{N}_{22} \end{pmatrix}, \; \Lambda = \begin{pmatrix} ½(\mathbf{N}_{11} + \mathbf{N}_{22} - \beta) & 0 \\ 0 & ½(\mathbf{N}_{11} + \mathbf{N}_{22} + \beta) \end{pmatrix},$$

$$\mathbf{V} = \begin{pmatrix} \alpha_1(\mathbf{N}_{11} - \mathbf{N}_{22} - \beta)/2\mathbf{N}_{21} & \alpha_2(\mathbf{N}_{11} - \mathbf{N}_{22} + \beta)/2\mathbf{N}_{21} \\ \alpha_1 & \alpha_2 \end{pmatrix},$$

$$\text{where } \beta = \sqrt{\mathbf{N}_{11}^2 - 2\mathbf{N}_{11}\mathbf{N}_{22} + 4\mathbf{N}_{12}\mathbf{N}_{21} + \mathbf{N}_{22}^2},$$

$$\alpha_1 = \left\{[(\mathbf{N}_{11} - \mathbf{N}_{22} - \beta)/2\mathbf{N}_{21}]^2 + 1\right\}^{-½}, \; \alpha_2 = \left\{[(\mathbf{N}_{11} - \mathbf{N}_{22} + \beta)/2\mathbf{N}_{21}]^2 + 1\right\}^{-½}.$$

Orthonormal matrix

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

If $\mathbf{V}$ is square and orthonormal then $\mathbf{V}^{\mathrm{T}} = \mathbf{V}^{-1}$.

Singular vectors and singular values

$$\mathbf{NV} = \mathbf{U}\Lambda, \qquad \mathbf{N}^{\mathrm{T}}\mathbf{U} = \mathbf{V}\Lambda, \qquad \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}_p, \qquad \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}_p,$$

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

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

$$\mathbf{N}^{\mathrm{T}}\mathbf{NV} = \mathbf{V}\Lambda, \qquad \mathbf{NN}^{\mathrm{T}}\mathbf{U} = \mathbf{U}\Lambda.$$

The trace of a matrix

Let $\mathbf{N}$ be the a square matrix. Its trace, tr ($\mathbf{N}$) is defined as

$$\operatorname{tr}(\mathbf{N}) = \sum_{i=1}^{n} \mathbf{N}_{ii}, \quad \mathbf{N} \in \mathbb{R}^{n \times n}.$$

The variance and standard deviation of a scalar

The variance is a statistical concept and requires a population of scalars, s^l, $1 \leqslant l \leqslant N$. The variance of s is approximated by the sample variance

$$\operatorname{var}(s) = \langle(s - \langle s\rangle)^2\rangle = \frac{1}{N-1}\sum_{l=1}^{N}(s^l - \langle s\rangle)^2,$$

$$\text{the sample mean } \langle s\rangle = \frac{1}{N}\sum_{l=1}^{N} s^l.$$

The standard deviation of s is

$$\sigma_s = \sqrt{\operatorname{var}(s)}.$$

N.B. The $\langle\bullet\rangle$ used as a mean in the above is different from the $\langle\bullet, \bullet\rangle$ involved in the definition of the scalar product between two vectors.

The covariance between two scalars

Consider a population of two scalars, s_{A}^l and s_{B}^l, $1 \leqslant l \leqslant N$. The sample covariance between s_{A} and s_{B} is

$$\operatorname{cov}(s_{\mathrm{A}}, s_{\mathrm{B}}) = \langle(s_{\mathrm{A}} - \langle s_{\mathrm{A}}\rangle)(s_{\mathrm{B}} - \langle s_{\mathrm{B}}\rangle)\rangle,$$

$$= \frac{1}{N-1} \sum_{l=1}^{N} (s_A^l - \langle s_A \rangle)(s_B^l - \langle s_B \rangle).$$

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

The correlation between two scalars

$$\mathrm{cor}(s_A, s_B) = \frac{\mathrm{cov}(s_A, s_B)}{\sigma_{s_A}\sigma_{s_B}}, \qquad -1 \leqslant \mathrm{cor}(s_A, s_B) \leqslant 1.$$

The covariance matrix between two vectors

Consider a population of two vectors, $\mathbf{v}_A^l$ and $\mathbf{v}_B^l$, $1 \leqslant l \leqslant N$. The sample covariance between $\mathbf{v}_A$ and $\mathbf{v}_B$ is

$$\mathrm{cov}(\mathbf{v}_A, \mathbf{v}_B) = \langle (\mathbf{v}_A - \langle \mathbf{v}_A \rangle)(\mathbf{v}_B - \langle \mathbf{v}_B \rangle)^T \rangle,$$

$$= \frac{1}{N-1} \sum_{l=1}^{N} (\mathbf{v}_A^l - \langle \mathbf{v}_A \rangle)(\mathbf{v}_B^l - \langle \mathbf{v}_B \rangle)^T,$$

$$\mathrm{cov}(\mathbf{v}_A, \mathbf{v}_B)_{ij} = \frac{1}{N-1} \sum_{l=1}^{N} ((\mathbf{v}_A^l)_i - \langle (\mathbf{v}_A)_i \rangle)((\mathbf{v}_A^l)_j - \langle (\mathbf{v}_A)_j \rangle),$$

$$\mathrm{cov}(\mathbf{v}_A, \mathbf{v}_B) \in \mathbb{R}^{m \times n}, \; \mathbf{v}_A^l \in \mathbb{R}^m, \; \mathbf{v}_B^l \in \mathbb{R}^n.$$

If $\mathbf{v}_A = \mathbf{v}_B (= \mathbf{v})$, then $\mathrm{cov}(\mathbf{v}, \mathbf{v})$ is the autocovariance matrix of $\mathbf{v}$ where $\mathbf{v} \in \mathbb{R}^n$, $\mathrm{cov}(\mathbf{v}, \mathbf{v}) \in \mathbb{R}^{n \times n}$. Diagonal elements are variances of each element of $\mathbf{v}$, i.e. $\mathrm{cov}(\mathbf{v}, \mathbf{v})_{ii} = \mathrm{var}(\mathbf{v}_i)$.

Matrix of standard deviations

$$\text{Let } \Sigma = \mathrm{diag}(\mathrm{var}^{1/2}(\mathbf{v}_1), \mathrm{var}^{1/2}(\mathbf{v}_2), \ldots),$$

$$\Sigma \in \mathbb{R}^{n \times n}, \; \mathbf{v} \in \mathbb{R}^n.$$

The correlation matrix between two vectors

$$\mathrm{cor}(\mathbf{v}, \mathbf{v}) = \Sigma^{-1} \mathrm{cov}(\mathbf{v}, \mathbf{v}) \Sigma^{-1},$$

$$\mathrm{cor}(\mathbf{v}, \mathbf{v})_{ij} = \frac{\mathrm{cov}(\mathbf{v}, \mathbf{v})_{ij}}{\mathrm{var}^{1/2}(\mathbf{v}_i)\, \mathrm{var}^{1/2}(\mathbf{v}_j)},$$

$$\mathrm{cov}(\mathbf{v}, \mathbf{v}) \in \mathbb{R}^{n \times n}, \quad \mathrm{cor}(\mathbf{v}, \mathbf{v}) \in \mathbb{R}^{n \times n}, \; \mathbf{v} \in \mathbb{R}^n.$$

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 ith column of $\mathbf{N}$ as a 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. The rank of a square matrix is also the number of non-zero eigenvalues.

The Fourier transform

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

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

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

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

The real-to-spectral space transform in higher dimensions (d-dimensions)

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

The spectral-to-real space transform in higher dimensions

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

The Fourier transform relies on the orthogonality relationships

$$\iiint_{\mathbf{x}} \exp(i\mathbf{k}\cdot\mathbf{x})\exp(-i\mathbf{k}'\cdot\mathbf{x})\,d\mathbf{x} = (2\pi)^d\,\delta(\mathbf{k}-\mathbf{k}'),$$

$$\iiint_{\mathbf{k}} \exp(i\mathbf{k}\cdot\mathbf{x})\exp(-i\mathbf{k}\cdot\mathbf{x}')\,d\mathbf{k} = (2\pi)^d\,\delta(\mathbf{x}-\mathbf{x}').$$

Convolution theorem

$$\int_{x'=-\infty}^{\infty} dx'\,g(x-x')\,f(x') \qquad \text{has F.T.} \qquad 2\pi\bar{g}(k)\,\bar{f}(k).$$

Fourier series

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

$$\bar{f}(k_i) = \frac{1}{\sqrt{N}}\sum_{j=1}^{N} f(x_j)\exp(-ik_ix_j),$$

$$f(x_j) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \bar{f}(k_i)\exp(ik_ix_j),$$

$$\sum_{j=1}^{N}\exp(ik_ix_j)\exp(-ik_{i'}x_j) = N\delta_{ii'},$$

$$\sum_{i=1}^{N}\exp(ik_ix_j)\exp(-ik_ix_{j'}) = N\delta_{jj'}.$$

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

$$\bar{\mathbf{f}} = \mathbf{F}\mathbf{f}, \qquad \mathbf{f} = \mathbf{F}^{\dagger}\bar{\mathbf{f}}, \qquad \mathbf{F}\mathbf{F}^{\dagger} = \mathbf{I}_N, \qquad \mathbf{F}^{\dagger}\mathbf{F} = \mathbf{I}_N,$$

where † means transform and complex conjugate. Matrix elements of $\mathbf{F}$ are

$$\mathbf{F}_{ij} = \frac{1}{\sqrt{N}}\exp(-ik_ix_j).$$

Lagrange multipliers

Problem: find the stationary point of $f(x_1, x_2, \ldots, x_N)$ subject to the constraints $g_m(x_1, x_2, \ldots, x_N) = 0$ for $1 \leqslant m \leqslant M$. This is a problem with N degrees of freedom and M constraints. The constrained variational problem is expressed as

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

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

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

where $\mathbf{x} = (x_1, x_2, \ldots, x_N)^{\mathrm{T}}$, and $\mathbf{G}_{mn} = \partial g_m/\partial x_n$.

The Sherman-Morrison-Woodbury formula

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

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

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

R.N.B. Nov 2010