

Math review

In our course, basically all geometries can be described by Cartesian coordinates in a Euclidean space. In the following, we give a very brief review of vector calculus.

Vectors, tensors

A **vector** \mathbf{u} can be written as

$$\mathbf{u} = u_x \hat{\mathbf{x}} + u_y \hat{\mathbf{y}} + u_z \hat{\mathbf{z}}$$

where $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ are the unit vectors in the x, y and z directions.

The **dot product** of two vectors is a scalar and defined as

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= u_x v_x + u_y v_y + u_z v_z \\ &= \sum_i u_i v_i = u_i v_i \end{aligned}$$

where we use the index notation i assuming it takes the values 1, 2 and 3 for the x, y and z components, respectively. Note that we also use the *summation convention*, i.e., repeated indices in a product are summed over.

A **second-order tensor** \mathbf{A} is a linear form that produces one vector from another, such as

$$\begin{aligned} \mathbf{u} &= \mathbf{A} \mathbf{v} \\ u_i &= \sum_j A_{ij} v_j = A_{ij} v_j \end{aligned}$$

where we sum again over indices $j = 1, 2, 3$.

The **double-dot product** of two second-order tensors is a scalar defined as

$$\mathbf{A} : \mathbf{B} = \sum_i \sum_j A_{ij} B_{ij} = A_{ij} B_{ij}$$

where we use the summation convention again over i and j . We will often use the double-dot product between a 4th-order tensor \mathbf{c} and a second-order tensor \mathbf{A} like

$$(\mathbf{c} : \mathbf{A})_{ij} = c_{ijkl} A_{kl}$$

summing over indices k and l , resulting in a second-order tensor.

The **dyadic product** or **tensor product** of two vectors is a tensor defined as

$$\begin{aligned} \mathbf{u} \mathbf{v} &= \mathbf{u} \otimes \mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & u_1 v_3 \\ u_2 v_1 & u_2 v_2 & u_2 v_3 \\ u_3 v_1 & u_3 v_2 & u_3 v_3 \end{bmatrix} \\ (\mathbf{u} \mathbf{v})_{ij} &= u_i v_j \end{aligned}$$

Note that the product is not commutative, i.e., the ordering of the vectors is sensitive and in general $\mathbf{u} \mathbf{v} \neq \mathbf{v} \mathbf{u}$.

Gradients

The **gradient** of a scalar field $\nabla\lambda$ is a vector field defined by the partial derivatives in x, y and z directions, i.e.,

$$\begin{aligned}\nabla\lambda &= \frac{\partial\lambda}{\partial x}\hat{\mathbf{x}} + \frac{\partial\lambda}{\partial y}\hat{\mathbf{y}} + \frac{\partial\lambda}{\partial z}\hat{\mathbf{z}} \\ (\nabla\lambda)_i &= \partial_i\lambda\end{aligned}$$

where we use the notation ∂_i as shorthand for $\partial/\partial x$, $\partial/\partial y$ and $\partial/\partial z$ for $i = 1, 2, 3$ respectively.

The **gradient** of a vector field $\nabla\mathbf{v}$ is a tensor field defined by

$$\begin{aligned}\mathbf{A} &= \nabla\mathbf{v} \\ A_{ij} &= \partial_i v_j\end{aligned}$$

The **divergence** of a vector field, written as $\nabla \cdot \mathbf{v}$ is a scalar field defined by

$$\begin{aligned}\nabla \cdot \mathbf{v} &= \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \\ &= \partial_i v_i\end{aligned}$$

using the summation convention.

The **divergence** of a second-order tensor field $\nabla \cdot \mathbf{A}$ is a vector field

$$(\nabla \cdot \mathbf{A})_j = \partial_i A_{ij}$$

where we again sum over indices $i = 1, 2, 3$.

Theorems

The **divergence theorem**, also known as **Gauss's theorem**, equates the volume integral of a vector field to the surface integral of the orthogonal component of the vector field, i.e.,

$$\int_V \nabla \cdot \mathbf{v} \, dV = \int_S \mathbf{v} \cdot \hat{\mathbf{n}} \, dS$$

where $\hat{\mathbf{n}}$ is the outward normal vector to the surface. Note that the same equality can be written for the divergence of a second-order tensor field.

The **Parseval's theorem** relates the time integral with the frequency integral of two functions, i.e.

$$\int f(t) g(t) \, dt = \frac{1}{2\pi} \int f(\omega) g^*(\omega) \, d\omega$$

where we use $*$ to denote the complex conjugate; $f(\omega)$ is the Fourier transform of $f(t)$, using (Dahlen & Tromp, 1998):

$$f(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad \text{forward transform}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega) e^{i\omega t} d\omega \quad \text{inverse transform}$$

Functions, Functionals and derivatives

The **chain rule** of the composition of two functions f and g , i.e., $f(g(x)) = (f \circ g)(x)$ gives the derivative defined by

$$\frac{d}{dx}(f(g(x))) = \frac{df}{dg} \frac{dg}{dx}$$

using Leibniz's notation.

The **derivation by parts** of the product of two functions $f(x)$ and $g(x)$ is defined by

$$\frac{d}{dx}(fg) = \frac{df}{dx}g + f\frac{dg}{dx}$$

As an example, the divergence of a vector field $\mathbf{v}(x)$ multiplied by a scalar $a(x)$ is a scalar defined by

$$\nabla \cdot (a \mathbf{v}) = \nabla a \cdot \mathbf{v} + a \nabla \cdot \mathbf{v}$$

The **integration by parts** of two functions $f(t)$ and $g(t)$ is defined by

$$\int_a^b f \partial_t g dt = f g \Big|_a^b - \int_a^b \partial_t f g dt$$

which can be seen from the derivation by parts $\frac{d}{dt}(fg) = \partial_t f g + f \partial_t g$. The term $fg \Big|_a^b$ on the right hand side is evaluated as $fg \Big|_a^b = f(b)g(b) - f(a)g(a)$.

A **functional** maps functions to scalars. In inverse problems, we use definitions of misfits between time series which are functionals. For example, using two traces with data $d(x, t)$ and synthetics $s(x, t)$ we can define the functional χ as:

$$\chi = \frac{1}{2} \int_0^T (s - d)^2 dt \quad \text{continuous form}$$

$$\tilde{\chi} = \sum_i (s_i - d_i)^2 \quad \text{discretized signals}$$

The **first variation** of a functional J is the Gâteaux derivative and defined as:

$$\delta J(y, h) = \lim_{\epsilon \rightarrow 0} \frac{J(y + \epsilon h) - J(y)}{\epsilon} = \frac{d}{d\epsilon} J(y + \epsilon h) \Big|_{\epsilon=0}$$

where h is an arbitrary function, used as the direction in which the functional varies.

We often use the **variation** of a function δf as the change of the function $f(x)$ when the parameter changes by a small amount:

Function	Variation
$f(x) = x^2$	$\delta f = 2x \delta x$
$f(g(x)) = a \frac{1}{g(x)}$	$\delta f = a \delta\left(\frac{1}{g}\right)$
$y = \sqrt{\frac{f(x)}{g(x)}} \Rightarrow y^2 = \frac{f(x)}{g(x)}$	$2y \delta y = \frac{1}{g} \delta f - \frac{f}{g^2} \delta g$ $\Leftrightarrow 2y^2 \frac{\delta y}{y} = \frac{f}{g} \frac{\delta f}{f} - \frac{f g}{g^2} \frac{\delta g}{g}$ $\Leftrightarrow 2 \frac{f}{g} \frac{\delta y}{y} = \frac{f}{g} \frac{\delta f}{f} - \frac{f}{g} \frac{\delta g}{g}$ $\Leftrightarrow \frac{\delta y}{y} = \frac{1}{2} \frac{\delta f}{f} - \frac{1}{2} \frac{\delta g}{g}$
$f(x) = \ln(x)$	$\delta f = \delta \ln(x) = \frac{\delta x}{x}$