AME50541: Finite Element Methods
Note 2: Mathematical Preliminaries

1 Introduction to vectors and tensors

A physical quantity completely defined by a single real number, e.g., temperature, density, mass, is a *scalar*. A *vector* is a directed line in space that represents physical quantities with a magnitude and direction. Geometric vectors define a quantity that is independent of the choice of basis, e.g., displacement and force. A collection of numbers such as $\boldsymbol{a} = (\text{your age}, \text{your height}, \text{your weight})$ is also commonly called a vector, but is not geometric. A tensor of order k is a geometric object (basis independent) that can be represented as k-dimensional array. Scalar and geometric vectors are tensors of order 0 and 1, respectively. For the remainder, we will solely consider geometric vectors and refer to them as vectors for brevity. Examples of higher order tensors are stress/strain (second-order) and the elasticity tensor (fourth-order).

Physical quantities are expressed in terms of their measurements in a coordinate system, defined by basis vectors $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_N\}$, (linearly independent) where $\boldsymbol{v} \in \mathbb{R}^N$ and N is the dimension of the space. We usually consider 2- or 3-dimensional systems ($N = 2, 3$). Vectors and tensors are represented by their components in a particular basis. Let $\boldsymbol{a}$ be a 3-dimensional vector and consider the basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$, then the $\boldsymbol{a} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + a_3\boldsymbol{v}_3$, where $\{a_1, a_2, a_3\}$ are the components of $\boldsymbol{a}$ in the basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$.

If the basis vectors are fixed, the coordinate system is Cartesian and rectangular if the basis vectors are orthogonal ($\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ if $i \neq j$). If the basis vectors are of unit length and mutually orthogonal, the basis is called orthonormal. The most common coordinate system is the rectangular Cartesian coordinate system defined by the canonical basis $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_N\}$, where $\boldsymbol{e}_i$ is zero except the ith entry which is 1.

2 Indicial notation

For convenience, brevity, and simplicity, we will make extensive use of *indicial notation* and Einstein summation convention that writes vector/tensor operations explicitly in terms of their components in an abbreviated form by eliminating summation symbols. Consider the representation of the vector $\boldsymbol{a}$ in the basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_N\}$ with components $\{a_1, \dots, a_N\}$,

$$\boldsymbol{a} = a_1\boldsymbol{v}_1 + \cdots + a_N\boldsymbol{v}_N = \sum_{i=1}^{N} a_i\boldsymbol{v}_i = a_i\boldsymbol{v}_i, \tag{1}$$

where the summation is implied in the last expression (summation convention). The summation convention says that whenever an index is repeated (twice) in the same term, then summation over its range is implied, unless otherwise stated. The index that is summed over is said to be the *dummy* index. Non-dummy indices are called *free* indices. Formally, the rules of indicial notation with Einstein summation convention are: 1) the same index may not appear more than twice in any term, 2) free indices on each term of the equation must agree, and 3) and free and dummy indices may be changed without altering the meaning of an expression, provided that rules 1 and 2 are not violated.

2.1 Kronecker delta

We define the Kronecker delta symbol δ_{ij} as

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{2}$$

The Kronecker delta also acts as a replacement operator that replaces a_i with a_j when multiplied by δ_{ij}, e.g., $a_i\delta_{ij} = a_1\delta_{1j} + \cdots + a_N\delta_{Nj} = a_j$.

2.2 Vector operations

Standard vector operations expressed using indicial notation

- Addition: $\boldsymbol{c} = \boldsymbol{a} + \boldsymbol{b} \rightarrow c_i = a_i + b_i$
- Dot product: $\lambda = \boldsymbol{a} \cdot \boldsymbol{b} = a_i b_i$
- Outer product: $\boldsymbol{S} = \boldsymbol{a} \otimes \boldsymbol{b} = \boldsymbol{a}\boldsymbol{b}^T \rightarrow S_{ij} = a_i b_j$
- Two norm: $\lambda = \|\boldsymbol{v}\|_2 = \sqrt{v_i v_i}$

2.3 Matrix operations

Standard matrix operations expressed using indicial notation

- Addition: $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} \rightarrow C_{ij} = A_{ij} + B_{ij}$
- Transpose: $\boldsymbol{A} = \boldsymbol{B}^T \rightarrow A_{ij} = B_{ji}$
- Double contraction: $\lambda = \boldsymbol{A} : \boldsymbol{B} = A_{ij} B_{ij}$
- Matrix-vector product:

$$\begin{aligned} \boldsymbol{c} = \boldsymbol{A}\boldsymbol{b} &\rightarrow c_i = A_{ij} b_j \\ \boldsymbol{c} = \boldsymbol{A}^T \boldsymbol{b} &\rightarrow c_i = A_{ji} b_j \end{aligned}$$

- Matrix-matrix product:

$$\begin{aligned} \boldsymbol{C} = \boldsymbol{A}\boldsymbol{B} &\rightarrow C_{ij} = A_{ik} B_{kj} \\ \boldsymbol{C} = \boldsymbol{A}^T \boldsymbol{B} &\rightarrow C_{ij} = A_{ki} B_{kj} \\ \boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}^T &\rightarrow C_{ij} = A_{ik} B_{jk} \\ \boldsymbol{C} = \boldsymbol{A}^T \boldsymbol{B}^T &\rightarrow C_{ij} = A_{ki} B_{jk} \end{aligned}$$

- Trace: $\lambda = \mathrm{tr}(\boldsymbol{A}) = A_{ii}$
- Frobenius norm: $\lambda = \|\boldsymbol{A}\|_F = \sqrt{\mathrm{tr}(\boldsymbol{A}\boldsymbol{A}^T)} = \sqrt{A_{ij} A_{ij}}$

3 Tensor calculus

A tensor field is a tensor quantity that varies in space and time. We will adopt the common convention as a shorthand notation for partial derivatives with respect to the coordinate variables in a Cartesian coordinate system. For a tensor of any order, all indices appearing *after* a comma indicate coordinates along which derivatives are taken; all indices that appearing before the comma have their usual meaning as indices of the tensor. The number of indices appear after the common determines the order of the derivative. For example, let ϕ be a scalar function, $\boldsymbol{A}$ be a second-order tensor, and $\boldsymbol{\mathcal{T}}$ be a tensor of order m, then

$$\phi_{,i} = \frac{\partial \phi}{\partial x_i} \qquad A_{ij,kl} = \frac{\partial^2 A_{ij}}{\partial x_k \partial x_l} \qquad T_{i_1 \cdots i_m, j} = \frac{\partial T_{i_1 \cdots i_m}}{\partial x_j}. \tag{3}$$

The **gradient** of a scalar is a vector whose components are its partial derivatives with respect to the coordinate variables in a Cartesian coordinate system:

$$\boldsymbol{g} = \nabla \phi = \begin{bmatrix} \phi_{,1} & \cdots & \phi_{,N} \end{bmatrix}^T \rightarrow g_i = \phi_{,i} \tag{4}$$

The definition generalizes to tensors of order m, where the gradient is a tensor of order $m+1$

$$\boldsymbol{G} = \nabla \boldsymbol{\mathcal{T}} \rightarrow G_{i_1 \ldots i_m j} = T_{i_1 \cdots i_m, j}. \tag{5}$$

The **divergence** of a vector is a scalar given by the contraction of the gradient: $h = \nabla \cdot \boldsymbol{v} = v_{i,i}$. The definition generalizes to tensors of order m, where the divergence is a tensor of order $m-1$

$$\boldsymbol{H} = \nabla \cdot \boldsymbol{\mathcal{T}} \rightarrow H_{i_1 \cdots i_{m-1}} = T_{i_1 \cdots i_m, i_m} \tag{6}$$

Finally, the **Laplacian** of a scalar function is a scalar defined as the divergence of the gradient

$$\Delta \phi = \nabla^2 \phi = \nabla \cdot \nabla \phi = \phi_{,ii}. \tag{7}$$

The definition generalizes to tensor of order m, where the Laplacian is a tensor of order m

$$\boldsymbol{F} = \Delta \boldsymbol{\mathcal{T}} = \nabla \cdot \nabla \boldsymbol{\mathcal{T}} \rightarrow F_{i_1 \cdots i_m} = T_{i_1 \cdots i_m, jj}. \tag{8}$$

4 Partial Differential Equations

A partial differential equation (PDE) is a relationship between an (unknown) function of several variables and its partial derivatives over a domain $\Omega \subset \mathbb{R}^N$ with boundary $\partial\Omega$. A function of several variables $(x_1, \ldots, x_N)$ is said to belong to be of class $C^m(\Omega)$ in the domain Ω if all its partial derivatives with respect to $(x_1, \ldots, x_N)$ of order up to and including m exist and are continuous in Ω. If the function only depends on one variable, it is called an ordinary differential equation. Let $u(x_1, \ldots, x_N, t)$ be an unknown function (dependent variable) with independent variables $x_1, \ldots, x_N, t$. The order of the PDE is determined by the highest derivative in the equation. A PDE is called *linear* if it is of first degree in all its field (dependent) variables and their partial derivatives.

Consider the second-order partial differential equation

$$A_{ij} u_{,ij} + (\text{lower-order terms}) = 0, \tag{9}$$

where $A_{ij} = A_{ij}(x_1, \ldots, x_N, t)$ are coefficients. The PDE is elliptic if the eigenvalues of the matrix with components A_{ij} are all positive or all negative. Elliptic PDEs always have smooth solutions, even if the initial or boundary conditions are non-smooth. If the matrix has all positive or all negative eigenvalues, except for one zero eigenvalue, the PDE is parabolic and represents a diffusion-like process, where information propagates throughout the domain instantaneously. If the matrix has one negative (positive) eigenvalue and the rest are positive (negative), the PDE is hyperbolic. The smoothness of the solution of hyperbolic PDEs depends on the smoothness of the initial data. For nonlinear hyperbolic PDEs, non-smooth features may develop even if the initial and boundary conditions are smooth.

5 Integral identities

The divergence theorem relates an integral over a domain Ω to one over its boundary. In one-dimension, this is the well-known fundamental theorem of calculus: let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be differentiable, then

$$\int_a^b \phi'(x)\, dx = [\phi]_a^b = \phi(b) - \phi(a). \tag{10}$$

In multiple dimensions, let Ω be an compact subset of $\mathbb{R}^n$ with piecewise smooth boundary $\partial\Omega$ and $\boldsymbol{F}$ be a continuously differentiable vector field on Ω, then

$$\int_\Omega \nabla \cdot \boldsymbol{F}\, dV = \int_{\partial\Omega} \boldsymbol{F} \cdot \boldsymbol{n}\, dS. \tag{11}$$

In the special case where $\boldsymbol{f}$ is a vector, $\boldsymbol{F}$ is a second-order tensor, and $\boldsymbol{\mathcal{T}}$ is tensor of order m, this reduces to

$$\begin{aligned} \int_\Omega f_{i,i}\, dV &= \int_{\partial\Omega} f_i n_i\, dS, \\ \int_\Omega F_{ij,j}\, dV &= \int_{\partial\Omega} F_{ij} n_j\, dS, \\ \int_\Omega T_{i_1 \cdots i_m, i_m}\, dV &= \int_{\partial\Omega} T_{i_1 \cdots i_m} n_{i_m}\, dS, \end{aligned} \tag{12}$$

respectively.

A particularly useful integral identity for our study of the finite element method is integration-by-parts, which follows directly from the divergence theorem, and allows one to move a derivative from one term in a product to another. Consider a vector-valued function $\boldsymbol{f}(\boldsymbol{x})$ and scalar function $w(\boldsymbol{x})$ over Ω, then the product rule of differentiation states

$$(f_i w)_{,i} = f_{i,i} w + f_i w_{,i}. \tag{13}$$

Integration over the domain and application of the divergence theorem yields

$$\begin{aligned} \int_\Omega w \nabla \cdot \boldsymbol{f} \, dV &= -\int_\Omega \boldsymbol{f} \cdot \nabla w \, dV + \int_{\partial\Omega} w \boldsymbol{f} \cdot \boldsymbol{n} \, dS \\ \int_\Omega f_{i,i} w \, dV &= -\int_\Omega f_i w_{,i} \, dV + \int_{\partial\Omega} w f_i n_i \, dS, \end{aligned} \tag{14}$$

where $\boldsymbol{n}$ is the outward normal to $\partial\Omega$. For the special case of a one-dimensional domain, this is known as integration-by-parts

$$\int_a^b w f' \, dx = -\int_a^b f w' \, dx + \left[w f \right]_a^b, \tag{15}$$

which can easily be derived from the product rule $\frac{d}{dx}(wf) = w'f + wf'$ and fundamental theorem of calculus. In the case where $\boldsymbol{f}(\boldsymbol{x}) = \nabla\phi(\boldsymbol{x})$, this relationship becomes

$$\begin{aligned} \int_\Omega w \Delta \phi \, dV &= -\int_\Omega \nabla\phi \cdot \nabla w \, dV + \int_{\partial\Omega} w \nabla\phi \cdot \boldsymbol{n} \, dS \\ \int_\Omega \phi_{,ii} w \, dV &= -\int_\Omega \phi_{,i} w_{,i} \, dV + \int_{\partial\Omega} w \phi_{,i} n_i \, dS. \end{aligned} \tag{16}$$