UCSB, Physics 129AL, Computational Physics
Lecture notes, Week 1

Zihang Wang (UCSB), zihangwang@ucsb.edu

January 9, 2025

Contents

1 Computation Complexity

1.1 Matrix and Tensor algebra in computation

Matrix algebra and tensor algebra are the basis of modern computations. Matrix operations such as multiplication, inversion, and eigenvalue analysis are standardized in numerical analysis and linear algebra, in computation, we sometimes refer higher-dimensional arrays as tensors. While matrices and tensors are interchangeable in computation, in physics tensors must obey strict transformation rules that preserve their geometric meaning under coordinate changes.

Those tensor transformation laws become important when modeling systems that involve conserved quantities or symmetries. In deep learning, for example, tensors not only serve as data structures but also could carry geometric information that encodes restrictions on how data transforms across layers.

In this chapter, I will establish foundational concepts on tensors. In 1.1.1, I will discuss vector spaces, the duality relation with covectors (one-forms), and the concept of inner products and metrics. Then, in 1.1.2, I will talk about general coordinate transformation laws, Jacobians, and the coordinate-free formalism. Finally, I will introduce you to the covariant derivative and Christoffel symbols.

1.1.1 Definition of vector and covector

Let's consider a m-dimensional vector space V with basis $\{\mathbf{e}_i\}$. Following the **Einstein summation** convention, we have,

$$\mathbf{n} = n^i \hat{\mathbf{e}}_i = \mathbf{n} = n^1 \hat{\mathbf{e}}_1 + n^2 \hat{\mathbf{e}}_2 + \cdots + n^i \hat{\mathbf{e}}_i \to \mathbf{n} = \begin{pmatrix} n^1 \\ n^2 \\ \vdots \\ n^i \end{pmatrix}. \tag{1}$$

However, the vector $\mathbf{n}$ should be independent from the underlying coordinate systems. When the coordinate system changes, the transformation of vector components n^i follows the inverse of the transformation applied to the basis vector $\hat{\mathbf{e}}_i$ such that it maintains the vector structure $\mathbf{n} = n'^i \hat{\mathbf{e}}_{i'} = n^i \hat{\mathbf{e}}_i$. We define the **contravariant vector**: the change in coordinates x'^i, x^i is opposite to the change in vector components,

$$n'^i = \frac{\partial x'^i}{\partial x^i} n^i, \qquad \hat{\mathbf{e}}_{i'} = \frac{\partial x^i}{\partial x'^i} \hat{\mathbf{e}}_i, \qquad \to n'^i \hat{\mathbf{e}}_{i'} = n^i \hat{\mathbf{e}}_i. \tag{2}$$

You can see that in this way, the coordinate transformation preserves the vector $\mathbf{n}$.

Let's make a simple example: in quantum mechanics, we have a general quantum state $\mathbf{n} \to |\Psi\rangle$, which is independent from the basis states:

$$|\Psi\rangle = \int dx \psi(x) |x\rangle = \int dp \phi(p) |p\rangle , \tag{3}$$

in this case, regardless which basis we use, either position or momentum, the physical quantum state $|\Psi\rangle$ remains unaltered.

The above invariance suggests a new type of vector space where its element $\boldsymbol{\omega}$ behaves in a covariant manner: the change in coordinates x'^i, x^i is aligned to the change in vector components,

$$\omega'_i = \frac{\partial x^i}{\partial x'^i} n_i, \qquad \hat{\mathbf{e}}_{i'} = \frac{\partial x'^i}{\partial x^i} \hat{\mathbf{e}}_i, \qquad \to n'_i \hat{\mathbf{e}}^{i'} = n_i \hat{\mathbf{e}}^i. \tag{4}$$

We call this vector space the **dual space** V^* of the vector space V, and the elements are **covectors**. It is a **covariant vector**: the change in coordinates x'^i, x^i is aligned with the change in vector components, A familiar example is the "bra" in quantum mechanics,

$$\langle\Psi| = \int dx \psi^*(x) \langle x| = \int dp \phi^*(p) \langle p| , \tag{5}$$

where $\psi^*(x), \phi^*(p)$ are complex conjugates.

More formally, let's consider a linear map $\boldsymbol{\omega} : V \to \mathbb{R}$ that lives in the **dual space** of V, denoted $\boldsymbol{\omega} \in V^*$, where the action of the linear map $\boldsymbol{\omega}$ on a vector

$\mathbf{n} \in V$, denoted $\boldsymbol{\omega}(\mathbf{n})$, produces a scalar (a **linear form**),

$$\boldsymbol{\omega}(\mathbf{n}) = n^i \omega_j \hat{\mathbf{e}}^j \hat{\mathbf{e}}_i = n^i \omega_j \delta_i^j = n^i \omega_i = \begin{pmatrix} \omega_x & \omega_y \cdots \end{pmatrix} \cdot \begin{pmatrix} n_x \\ n_y \\ \vdots \end{pmatrix} \in \mathbb{R}, \tag{6}$$

such that the linear map $\hat{\mathbf{e}}^j$ brings the vector basis $\hat{\mathbf{e}}_j$ to Kronecker delta δ_i^j $\hat{\mathbf{e}}^j \hat{\mathbf{e}}_i = \delta_i^j$. This is the duality relationship, which should be very familiar to you as the inner product. For example, "overlap" between quantum states, $\langle \Psi | \Phi \rangle$, is independent from any underlying coordinate system and basis transformations. We should note that while tshe vector basis and covector basis transform differently from their component, they remain in their respective vector spaces.

Let's consider a position-dependent linear map $\mathcal{O} : V \to V^*$ that generates one-to-one transformations (isomorphism) between a vector space and its dual space,

$$\hat{\mathbf{e}}^i = \mathcal{O}(\mathbf{r}) \hat{\mathbf{e}}_j = \sum_j g(\mathbf{r}, i, j) \hat{\mathbf{e}}_j = g^{ij} \hat{\mathbf{e}}_j, \quad \hat{\mathbf{e}}^k \hat{\mathbf{e}}^i = g^{ik}, \tag{7}$$

where $g(\mathbf{r}) \in \mathbb{R}$ is a position dependent variable.

$$\mathcal{O}^{-1}(\mathbf{r}) \hat{\mathbf{e}}^i = \hat{\mathbf{e}}_j = \sum_i \frac{1}{g(\mathbf{r}, i, j)} \hat{\mathbf{e}}^i = g_{ij} \hat{\mathbf{e}}^i, \quad \hat{\mathbf{e}}_k \hat{\mathbf{e}}_j = g_{kj}. \tag{8}$$

We should note that the notation $\hat{\mathbf{e}}_k \hat{\mathbf{e}}_j$: it does not produce a scalar, and instead, it is a linear map that transforms a vector to a covector, or it is a **bilinear form** that takes 2 vectors (covectors) as inputs and produce a output. Therefore, it is not a coordinate invariant measure.

It is convenient if we only want to work within a single vector space, and using Eq. 7

$$\boldsymbol{\omega}(\mathbf{n}) = n^i \hat{\mathbf{e}}_i \omega_j \hat{\mathbf{e}}^j = n^i \omega_j \hat{\mathbf{e}}_i g^{jk} \hat{\mathbf{e}}_k = n^i \omega^k \hat{\mathbf{e}}_i \hat{\mathbf{e}}_k = n^i \omega^k g_{ik} = \boldsymbol{\omega} \cdot \mathbf{n}, \tag{9}$$

such that $\omega^k = g^{jk} \omega_j$ and $\boldsymbol{\omega}, \mathbf{n} \in V$. There is no free lunch: we have to deal with a new object with a new name, metric g_{ik}. The metric is symmetric. If we are dealing with flux (or current), anti-symmetric tensor is needed, e.g. Levi-Civita tensor, $+1$ for even permutations of the indices.

You might recall that it is the definition of an inner product: an coordinate independent measure that takes two vectors as inputs and produce one scalar as the output. Again, it is a bilinear form, and the covector is not a linear transformation (operator). In a **flat space** with a Cartesian coordinate system, the metric is a Kronecker delta $g_{ik} = \delta_{ik}$, but on a general surface (2-dimensional manifold), parameterized by (u, v), or different coordinate systems, the metric tensor can be calculated via the following formula,

$$g_{ik} = \frac{\partial \mathbf{r}}{\partial i} \cdot \frac{\partial \mathbf{r}}{\partial k}, \quad i, k \in \{u, v\}, \tag{10}$$

where the position vector is $\mathbf{r}(u, v)$ is a point in the ambient space (embedding a $\mathbb{R}^2$ surface in $\mathbb{R}^3$). By symmetry, we have $g_{uv} = g_{vu}$, and this tensor describes the intrinsic geometry of the surface. For example, the metric tensor gives the notion of unit length on the surface,

$$g(u,v) = \begin{pmatrix} du & dv \end{pmatrix} \begin{pmatrix} \frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial u} & \frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v} \\ \frac{\partial \mathbf{r}}{\partial v} \cdot \frac{\partial \mathbf{r}}{\partial u} & \frac{\partial \mathbf{r}}{\partial v} \cdot \frac{\partial \mathbf{r}}{\partial v} \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}. \tag{11}$$

1.1.2 Coordinate Free Formalism

When you're dealing with vectors, covectors, and tensors, you focus on their geometric properties rather than the specific coordinate system you're working in. The transformation of the components is implicitly understood through the tensor transformation rules, and you don't necessarily need to spell out how the basis vectors change each time during the transformation. Vectors and covectors are invariant under coordinate transformations $x^i \to x^{i'}$,

$$n^{i'} = \frac{\partial x^{i'}}{\partial x^i} n^i, \text{ and } \omega_{i'} = \frac{\partial x^i}{\partial x^{i'}} \omega_{i'}, \qquad \hat{\mathbf{e}}^{i'} = \frac{\partial x^{i'}}{\partial x^i} \hat{\mathbf{e}}^i, \text{ and } \hat{\mathbf{e}}_{i'} = \frac{\partial x^i}{\partial x^{i'}} \hat{\mathbf{e}}_i, \tag{12}$$

where the index are implicitly summed as previously mentioned. If we know the transformation law of components, we immediately know the transformation law of corresponding basis. Therefore, from now on, we will only focus on the transformations of components until we get to the parallel transport.

The above coodinate transformation defines the **Jacobian matrix**: it contains partial derivatives that describes how coordinates transform between two coordinate systems, x^i and $x^{i'}$,

$$J_i^{i'} = \frac{\partial x^{i'}}{\partial x^i}, \qquad (J^{-1})_{i'}^i = \frac{\partial x^i}{\partial x^{i'}}. \tag{13}$$

We should note that the Jacobian is **not** a bilinear form, it takes one vector and one covector as inputs, and produces a scalar (a (1,1) tensor).

For a vector n^i and a covector ω^i, its transformation law is given by:

$$n^{i'} = \frac{\partial x^{i'}}{\partial x^i} n^i, \qquad \omega_{i'} = \frac{\partial x^i}{\partial x^{i'}} \omega_i. \tag{14}$$

The Jacobian captures the rate of change of new coordinates with respect to the old coordinates, and its inverse is used to transform covectors. This matrix is critical in various computational applications, particularly the concept of the back propagation.

Let's look at a example, in the transformation from Cartesian coordinates (x, y, z) to spherical coordinates (r, θ, ϕ), the coordinate transformation is given by:

$$x = r \sin\theta \cos\phi, \quad y = r \sin\theta \sin\phi, \quad z = r\cos\theta. \tag{15}$$

To find how the vector components transform between the coordinate systems, we calculate the Jacobian matrix J, which is composed of partial derivatives of

x, y, z with respect to r, θ, ϕ:

$$J = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{pmatrix}, \tag{16}$$

where each of the partial derivatives is computed as follows:

$$\frac{\partial x}{\partial r} = \sin\theta\cos\phi, \quad \frac{\partial x}{\partial \theta} = r\cos\theta\cos\phi, \quad \frac{\partial x}{\partial \phi} = -r\sin\theta\sin\phi,$$

$$\frac{\partial y}{\partial r} = \sin\theta\sin\phi, \quad \frac{\partial y}{\partial \theta} = r\cos\theta\sin\phi, \quad \frac{\partial y}{\partial \phi} = r\sin\theta\cos\phi,$$

$$\frac{\partial z}{\partial r} = \cos\theta, \quad \frac{\partial z}{\partial \theta} = -r\sin\theta, \quad \frac{\partial z}{\partial \phi} = 0.$$

Therefore, the Jacobian matrix becomes:

$$J = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}. \tag{17}$$

For a vector $v^i = (v^x, v^y, v^z)$ in Cartesian coordinates, its components in spherical coordinates $v^{i'} = (v^r, v^\theta, v^\phi)$ are obtained by multiplying the Jacobian matrix with the vector components:

$$v^{i'} = J^{i'}_i v^i. \tag{18}$$

Similarly, for a covector $\omega_i = (\omega_x, \omega_y, \omega_z)$, the components in spherical coordinates are given by the inverse Jacobian matrix:

$$\omega_{i'} = (J^{-1})^i_{i'}\omega_i. \tag{19}$$

1.1.3 Gradient, covariant derivative and Christoffel symbols

Let's consider the gradient of a vector,

$$\partial_j \mathbf{n} = (\partial_j n^i)\hat{\mathbf{e}}_i + n^i \partial_j \hat{\mathbf{e}}_i, \tag{20}$$

where we use the product rule. The first term $(\partial_j n^i)$ captures the magnitude variation in the same direction $\hat{\mathbf{e}}_i$, and the second term $n^i \partial_j \hat{\mathbf{e}}_i$ captures the change in the basis state. The partial derivative transforms as a covariant vector,

$$\partial_{i'} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial}{\partial x^i} = (J^{-1})^i_{i'} \partial_i. \tag{21}$$

For example, consider a vector field $\mathbf{n} = n^r \hat{\mathbf{e}}_r + n^\theta \hat{\mathbf{e}}_\theta + n^\phi \hat{\mathbf{e}}_\phi$ in spherical coordinates (r, θ, ϕ). The partial derivative with respect to x in Cartesian coordinates can be expressed in terms of the spherical coordinate derivatives using the inverse Jacobian:

$$\partial_x = \frac{\partial r}{\partial x}\partial_r + \frac{\partial \theta}{\partial x}\partial_\theta + \frac{\partial \phi}{\partial x}\partial_\phi. \tag{22}$$

Thus, for the derivative of the vector field $\mathbf{n}$, we have:

$$\partial_x \mathbf{n} = \frac{\partial r}{\partial x}\partial_r(n^r \hat{\mathbf{e}}_r) + \frac{\partial \theta}{\partial x}\partial_\theta(n^\theta \hat{\mathbf{e}}_\theta) + \frac{\partial \phi}{\partial x}\partial_\phi(n^\phi \hat{\mathbf{e}}_\phi). \tag{23}$$

The term $\partial_j \hat{\mathbf{e}}_i$ can be expressed via Christoffel symbols $\partial_j \hat{\mathbf{e}}_k = \Gamma^i_{jk}\hat{\mathbf{e}}_i$,

$$\partial_j \mathbf{n} = \partial_j n^i \hat{\mathbf{e}}_i + n^i \partial_j \hat{\mathbf{e}}_i = \partial_j n^i \hat{\mathbf{e}}_i + n^k \Gamma^i_{jk}\hat{\mathbf{e}}_i = (\partial_j n^i + n^k \Gamma^i_{jk})\hat{\mathbf{e}}_i = (\nabla_j n^i)\hat{\mathbf{e}}_i. \tag{24}$$

We call $\nabla_j n^i$ the **covariant derivative**. The Christoffel symbols are derived by permutations of metric compatibility condition: the covariant derivative of the metric vanishes $\nabla_j g_{ik} = 0$. This is trivially follow from the definition of the metric tensor previously: Recall that the metric tensor is a bilinear form $\hat{\mathbf{e}}_i \hat{\mathbf{e}}_k = g_{ik}$. Using the product rule and the

$$\begin{aligned}
\partial_j g_{ki} - \partial_j(\hat{\mathbf{e}}_k \cdot \hat{\mathbf{e}}_i) = \partial_j g_{ki} - (\partial_j \hat{\mathbf{e}}_k)\cdot \hat{\mathbf{e}}_i - \hat{\mathbf{e}}_k \cdot (\partial_j \hat{\mathbf{e}}_i) = 0 \\
0 = \partial_j g_{ki} - \Gamma^m_{jk}\hat{\mathbf{e}}_m \cdot \hat{\mathbf{e}}_i - \hat{\mathbf{e}}_k \cdot \hat{\mathbf{e}}_m \Gamma^m_{ji} \\
0 = \partial_j g_{ki} - \Gamma^m_{jk} g_{mi} - \Gamma^m_{ji} g_{km} \\
0 = \nabla_j g_{ki},
\end{aligned} \tag{25}$$

where we also get the definition of covariant derivative of a rank-2 tensor,

$$\nabla_j T_{ki} = \partial_j g_{ki} - \Gamma^m_{jk} T_{mi} - \Gamma^m_{ji} T_{km}, \tag{26}$$

and it has nice property: the covariant derivative of a tensor is still a tensor, unlike the partial derivatives. We will talk more about it soon.

Since the metric is symmetric ($g_{ik} = g_{ki}$), the partial derivatives $\nabla_k g_{ij}, \nabla_j g_{ki}, \nabla_i g_{jk}$ can be cyclically permuted. Let's write down three versions of the metric compatibility equation by permuting indices:

For j, k, i,

$$\nabla_j g_{ki} = \partial_j g_{ki} - \Gamma^m_{jk} g_{mi} - \Gamma^m_{ji} g_{km} = 0. \tag{27}$$

For k, i, j:

$$\nabla_k g_{ij} = \partial_k g_{ij} - \Gamma^m_{ki} g_{mj} - \Gamma^m_{kj} g_{im} = 0. \tag{28}$$

For i, j, k:

$$\nabla_i g_{jk} = \partial_i g_{jk} - \Gamma^m_{ij} g_{mk} - \Gamma^m_{ik} g_{jm} = 0. \tag{29}$$

Add the first and third equations and subtract the first,

$$(\partial_i g_{jk} + \partial_j g_{ki}) - (\partial_k g_{ij}) = (\Gamma^m_{ji} + \Gamma^m_{ij}) g_{km} + \Gamma^m_{jk} g_{mi} + \Gamma^m_{ik} g_{jm} - \Gamma^m_{ki} g_{mj} - \Gamma^m_{kj} g_{im}. \tag{30}$$

Using the symmetry of Christoffel symbols ($\Gamma^k_{ij} = \Gamma^k_{ji}$), the above expression simplifies to,

$$(\partial_i g_{jk} + \partial_j g_{ki}) - (\partial_k g_{ij}) = 2\Gamma^m_{ij} g_{km}. \tag{31}$$

To isolate Γ^k_{ij}, we can multiply both sides by the inverse metric and divide by 2 to get the expression for the Christoffel symbols Γ^m_{ij},

$$\Gamma^m_{ij} = \frac{1}{2} g^{km} \left(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij} \right). \tag{32}$$

With the Cartesian coordinate system, the connection Γ^m_{ij} vanishes, so that the coefficient does not change. However, if the space is non-flat or the coordinate system is **curvilinear** (with linear coordinate transformation), we must incorporate the Christoffel symbols in derivatives.

One known example is the geometric phase, where a vector (wavefuntion) travels on a curved surface accumulates an extra phase, resulting from non-trivial geometric connections. We will discuss it when dealing with parallel transport.

Let's look at an example. In spherical coordinates (r, θ, ϕ), the metric tensor g_{ij} is given by,

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix}, \tag{33}$$

and the inverse metric tensor g^{ij} is,

$$g^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2\theta} \end{pmatrix}. \tag{34}$$

Using the above information, we can calculate the nonzero Christoffel symbols for spherical coordinates using the derived formula,

$$\Gamma^m_{ij} = \frac{1}{2} g^{km} \left(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij} \right). \tag{35}$$

For example, we calculate the connection $\Gamma^r_{\theta\theta}$,

$$\begin{aligned} \Gamma^m_{ij} = \Gamma^r_{\theta\theta} &= \frac{1}{2} g^{kr} \left(\partial_\theta g_{\theta k} + \partial_\theta g_{k\theta} - \partial_k g_{\theta\theta} \right) \\ &= \frac{1}{2} \cdot (0 + 0 - 1 \cdot (2r)) = -r. \end{aligned} \tag{36}$$

We should note that Γ^m_{ij} is not a tensor: it does not transform linearly with the Jacobian.

Following the above discussion, let's consider the following: Suppose you are in $\mathbb{R}^3$, and you decide to use spherical coordinates to describe a flat surface $\gamma(u, v)$. Due to the reference coordinate system, the metric is not trivial. Are there coordinate independent measures that capture intrinsic properties of the parametrized surface?