

Notes for 3510

`This is a big reference sheet/summary for me to look back to while I'm taking this class. Meant to go over the major points and key ideas of the course, and some fun stuff :)

Chapter 1: Vectors and Matrices

1.1 Vectors in \mathbb{R}^n

// TODO

1.2 Dot Product

// TODO

1.3 Subspaces of \mathbb{R}^n

// TODO

1.4 Linear Transformations and Matrix Algebra

// TODO

1.5 Introduction to Determinants and the Cross Product

// TODO

Chapter 2: Functions, Limits, and Continuity

2.1 Scalar- and Vector-Valued Functions

// TODO

2.2 A Bit of Topology in \mathbb{R}^n

// TODO

2.3 Limits and Continuity

// TODO

Chapter 3: The Derivative

3.1 Partial Derivatives and Directional Derivatives

// TODO

3.2 Differentiability

// TODO

3.3 Differentiation Rules

// TODO

3.4 The Gradient

// TODO

3.5 Curves

// TODO

3.6 Higher-Order Partial Derivatives

// TODO

Chapter 4: Implicit and Explicit Solutions of Linear Systems

4.1 Gaussian Elimination and the Theory of Linear Systems

// TODO

4.2 Elementary Matrices and Calculating Inverse Matrices

// TODO

4.3 Linear Independence, Basis, and Dimension

// TODO

4.4 The Four Fundamental Subspaces

// TODO

4.5 The Nonlinear Case: Introduction to Manifolds

// TODO

Chapter 5: Extremum Problems

5.1 Compactness and the Maximum Value Theorem

// TODO

5.2 Maximum/Minimum Problems

// TODO

5.3 Quadratic Forms and the Second Derivative Test

// TODO

5.4 Lagrange Multipliers

// TODO

5.5 Projections, Least Squares, and Inner Product Spaces

// TODO

Chapter 6: Solving Nonlinear Problems

6.1 The Contraction Mapping Principle

Convergence of Vector Sums

If you have a convergent sequence of vectors $\{\mathbf{a}_k\}$, then this convergent sequence of vectors will converge as long as the norms of the vectors converge

$$\sum_{k=1}^{\infty} \|\mathbf{a}_k\| \text{ converges} \implies \sum_{k=1}^{\infty} \mathbf{a}_k \text{ converges}$$

Contraction Mappings

A contraction mapping is a function that always shrinks the distance between two points.

For $X \subset \mathbb{R}^n$, $f : X \rightarrow X$, f is a contraction mapping if $\exists c \in (0, 1)$ such that

$$\|f(x) - f(y)\| \leq c\|x - y\|$$

Contraction mappings are always continuous, and have at most one fixed point. If X is closed, we can specify that such a point does exist. This is called the contraction mapping principle. The contraction mapping principle states that if you call a contraction mapping f enough times, then it eventually converges on some point, called the fixed point of f .

if X is closed, then $\exists x \in X$ so that $f(x) = x$

The Mean Value Inequality

The MVT in single-var calc tells us that, on a continuous curve, there is some point where the tangent line at that point has the same slope as the secant line between the two endpoints. i.e. that $\frac{f(b) - f(a)}{b - a} = f'(x)$ for some $x \in (a, b)$. If we rearrange, however, this result tells us how much a function stretches or shrinks two points based on the slope of the function. This idea is how we generalize the MVT to \mathbb{R}^n

If $U \subset \mathbb{R}^n$ is open, $f : U \rightarrow \mathbb{R}^m$, and $a, b \in U$ with a \mathcal{C}^1 path in U which joins them, then

$$\|f(b) - f(a)\| \leq \left(\max_{x \in [a, b]} \|Df(x)\| \right) \|b - a\|$$

This gives us a really useful condition which we can use to prove that certain functions are contraction mappings. (if $0 < \max_{x \in [a, b]} \|Df(x)\| < 1$).

Matrix Theorems

Generalization of a geometric series. If $A \in \mathcal{M}_{n \times n}$ and $\|A\|_{\text{op}} < 1$, then

$$\sum_{i=0}^{\infty} A^i = (I - A)^{-1}$$

The general linear group $GL(n)$ of invertible matrices is an open subset of $\mathcal{M}_{n \times n}$, and $f : GL(n) \rightarrow GL(n)$ given by $f(A) = A^{-1}$ is smooth.

Since $\mathcal{M}_{m \times n}$ is isomorphic to \mathbb{R}^{nm} , we can take derivative maps of matrix-valued functions. For example, consider $f : \mathcal{M}_{n \times n} \rightarrow \mathcal{M}_{n \times n}$ given by $f(A) = A^2$. We can use the limit definition similarly to vectors (with respect to the Frobenius norm) and show $[Df(A)]B = AB + BA$. What this says, is that if we associate a linear map to $[Df(A)] : \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$, then that linear map takes a $B \in \mathbb{R}^{n^2}$ to $AB + BA$ (a sort of matrix power rule, which unfortunately doesn't simplify since multiplication is non-commutative).

The symmetric matrices where $A^T = A$ form an $\frac{1}{2}n(n+1)$ -dimensional subspace. The orthogonal group $O(n) \subset \mathcal{M}_{n \times n} = \{\text{set of matrices where } A^T A = I\}$ is a $\frac{1}{2}n(n-1)$ -dimensional submanifold, and $T_I O(n)$ (tangent plane at the identity) is the set of skew-symmetric matrices $\text{Skew}(n) = \{\text{set of matrices where } A^T = -A\}$

6.2 The Inverse and Implicit Function Theorems

Inverse Function Theorem

The Inv. FT is an incredibly powerful theorem which tells us under what conditions a certain function f has a local inverse, and what the derivative of that inverse looks like. To have an inverse, the derivative map must be locally invertible, and the inverse of that derivative map is the derivative of the inverse.

More formally it says if $U \subset \mathbb{R}^n$ is open, $f : U \rightarrow \mathbb{R}^n$ is \mathcal{C}^1 , $\mathbf{x}_0 \in U$ and $Df(\mathbf{x}_0)$ is invertible, then \exists two open neighborhoods, $V \subset U$ of \mathbf{x}_0 and W of $\mathbf{y}_0 := f(\mathbf{x}_0)$, and a \mathcal{C}^1 function between them $g : W \rightarrow V$ so that

$$\forall \mathbf{x} \in V, g(f(\mathbf{x})) = \mathbf{x}$$

$$\forall \mathbf{y} \in W, f(g(\mathbf{y})) = \mathbf{y}$$

Furthermore, $Dg(f(x)) = (Df(x))^{-1}$

Implicit Function Theorem

The Imp. FT, a consequence of the Inv. FT, is another incredibly powerful result. It tells us that a series of m implicit equations of n variables can be written locally as m explicit functions of $n - m$ variables (as long as the derivative map of the output variables is invertible). It is like a non-linear generalization of pivot and free variables. This is also incredibly useful for our future study of manifolds; it links the implicit description to the explicit description and allows us to think of manifolds locally as spaces where we can perform calculus like it's linear algebra.

More formally, for the Imp. FT we need $n > m$, and a \mathcal{C}^1 function

$\mathbf{F} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ where U is open. Then, writing a vector $\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \in \mathbb{R}^n$ with

$\mathbf{x} \in \mathbb{R}^{n-m}$ and $\mathbf{y} \in \mathbb{R}^m$, suppose $\mathbf{F} \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix} = \vec{0}$ describes a series of m implicit equations and $\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix} \in \mathcal{M}_{m \times m}$ is invertible.

Then, \exists two neighborhoods, V of \mathbf{x}_0 and W of \mathbf{y}_0 and a \mathcal{C}^1 function between them $\varphi : V \rightarrow W$ so that for all $\mathbf{x} \in V$ and $\mathbf{y} \in W$,

$$\mathbf{F} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \vec{0} \iff \mathbf{y} = \varphi(\mathbf{x})$$

Furthermore, $D\varphi(\mathbf{x}) = - \left(\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \begin{pmatrix} \mathbf{x} \\ \varphi(\mathbf{x}) \end{pmatrix} \right)^{-1} \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}} \begin{pmatrix} \mathbf{x} \\ \varphi(\mathbf{x}) \end{pmatrix} \right)$.

It helps to keep track of dimensions. $\frac{\partial \mathbf{F}}{\partial \mathbf{y}}$ is an $m \times m$ matrix (differentiating m functions w.r.t m variables), $\frac{\partial \mathbf{F}}{\partial \mathbf{x}}$ is an $m \times (n - m)$ matrix (differentiating m functions w.r.t. $(n-m)$ variables), so $D\varphi(\mathbf{x}) \in \mathcal{M}_{m \times (n-m)}$.

6.3 Manifolds Revisited

Manifolds are essentially special surfaces which are smooth and are very nice to deal with when doing calculus.

There are three equivalent definitions for a k -dimensional manifold M .

1. Explicit: a graph of a smooth function
2. Implicit: solutions to an implicit set of equations
3. Parametric: parameterized by an injective function that never has velocity 0

more formally

1. For any $\mathbf{p} \in M$, \exists a neighborhood $W \subset \mathbb{R}^n$ of \mathbf{p} so that $W \cap M$ is the graph of a smooth function $\mathbf{f} : V \subset \mathbb{R}^k \rightarrow \mathbb{R}^{n-k}$, where V is open.
Remember that $\text{graph}\{\mathbf{f}(\mathbf{x})\} = \left\{ \begin{bmatrix} \mathbf{x} \\ \mathbf{f}(\mathbf{x}) \end{bmatrix} : \mathbf{x} \in \mathbb{R}^k \right\}$, but that the order of the x_i 's doesn't matter.
2. For any $\mathbf{p} \in M$, \exists a neighborhood $W \subset \mathbb{R}^n$ of \mathbf{p} and a smooth function $\mathbf{F} : W \rightarrow \mathbb{R}^{n-k}$ so that the preimage (level set) $\mathbf{F}^{-1}(\vec{0}) = W \cap M$ and $\text{rank}\{D\mathbf{F}(\mathbf{x})\} = n - k$ for every $\mathbf{x} \in W \cap M$.
3. For any $\mathbf{p} \in M$, \exists a neighborhood $W \subset \mathbb{R}^n$ of \mathbf{p} so that $W \cap M$ is the image of a smooth function $\mathbf{g} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$, where U is open. \mathbf{g} must be injective, have $\text{rank}\{D\mathbf{g}(\mathbf{u})\} = k$ for all $\mathbf{u} \in U$, and have $\mathbf{g}^{-1} : W \cap M \rightarrow U$ be continuous.

We can also find the tangent planes at \mathbf{p} to each respective manifold ($T_{\mathbf{p}}M$) as follows

1. If M is locally at \mathbf{p} a graph of \mathbf{f} with $\mathbf{p} = \begin{bmatrix} \mathbf{a} \\ \mathbf{f}(\mathbf{a}) \end{bmatrix}$, then
 $T_{\mathbf{p}}M = \text{graph}\{D\mathbf{f}(\mathbf{a})\}$.
2. If M is locally a preimage of \mathbf{F} , then $T_{\mathbf{p}}M = \ker\{D\mathbf{F}(\mathbf{p})\}$.
3. If M is locally parameterized by \mathbf{g} with $\mathbf{p} = \mathbf{g}(\mathbf{a})$, then
 $T_{\mathbf{p}}M = \text{im}\{D\mathbf{g}(\mathbf{a})\}$

Chapter 7: Integration

7.1 Multiple Integrals

Defining Integrability (in higher dimensions)

To extend our notion of integration into multiple dimensions, we will be working with something called the Darboux integral. It is much like our notion of 2-D area under the curve, but generalizes to volumes of surfaces/hypersurfaces.

Given a rectangle $R \subset \mathbb{R}^n$, you can slice up the rectangle into subrectangles into a partition $\mathcal{P} := \mathcal{P}_1 \times \dots \times \mathcal{P}_n$ (each \mathcal{P}_i for $1 \leq i \leq n$ is like a list of points where to slice the rectangle in the corresponding dimension), where each $R_{i,j}$ of \mathcal{P} is the i - j th rectangle of R . Although for \mathbb{R}^n it should be more like the (i, j, k, \dots, n) -th rectangle of R for all the n coordinates needed to specify which subrectangle it is. For ease of notation I'll stick with $R_{i,j}$.

Recall that for a bounded function $f : R \rightarrow \mathbb{R}$, $\sup_{\mathbf{x} \in S} f(\mathbf{x})$ is the "least upper bound" and $\inf_{\mathbf{x} \in S} f(\mathbf{x})$ is the "greatest lower bound" of f on S . That is to say, \sup is like a maximum of f and \inf is like a minimum of f (although the function might never reach their respective values, only approach them).

We then define two possible sums on a partition \mathcal{P} , the upper and lower sums. First, we define a sort of minimum/maximum for a subrectangle $R_{i,j}$ to be $m_{i,j} := \inf_{\mathbf{x} \in R_{i,j}} f(\mathbf{x})$ and $M_{i,j} := \sup_{\mathbf{x} \in R_{i,j}} f(\mathbf{x})$, respectively. Then, the lower

sum of a partition is $L(f, \mathcal{P}) := \sum_{i,j} m_{i,j} \text{area}(R_{i,j})$ and the upper sum is

$$U(f, \mathcal{P}) := \sum_{i,j} M_{i,j} \text{area}(R_{i,j}).$$

Finally, we define integrability. f is integrable on R if for all partitions \mathcal{P} of R there is a *unique* number I satisfying $L(f, \mathcal{P}) \leq I \leq U(f, \mathcal{P})$. If f is integrable

we say $\int_R f dV := I$.

Proving Integrability

A nice way to prove integrability is the convenient criterion. If f is bounded on R , f is integrable on R iff $\forall \epsilon > 0$, \exists a partition \mathcal{P} of R such that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \epsilon$.

If f is continuous, and only discontinuous on a set of volume 0, then f is integrable. X having volume 0 means that \exists finite R_1, \dots, R_s so that $X \subset R_1 \cup \dots \cup R_s$ and $\text{vol}(R_1) + \dots + \text{vol}(R_s) < \epsilon$ for any $\epsilon > 0$.

Regions

A subset Ω is a region if it is the closure of a bounded open subset of \mathbb{R}^n and its boundary has volume 0. In other words, if $\Omega = \bar{U}$ for some open U , and $\partial\Omega$ has volume 0.

To integrate f over a region Ω , we pick a rectangle R with $\Omega \subset R$. Then, we

define $\tilde{f} = \begin{cases} f(\mathbf{x}) & \mathbf{x} \in \Omega \\ 0 & \text{otherwise} \end{cases}$

Then, $\int_{\Omega} f dV = \int_R \tilde{f} dV$

The volume of a region Ω is given by $\text{vol}(\Omega) = \int_{\Omega} 1 dV$

Properties of Integrals

1. Integration is linear

$$1. \int_{\Omega} (f + g) dV = \int_{\Omega} f dV + \int_{\Omega} g dV$$

$$2. \int_{\Omega} \alpha f dV = \alpha \int_{\Omega} f dV$$

$$2. \int_{A \cup B} f dV = \int_A f dV + \int_B f dV$$

$$3. \text{ If } f \leq g \text{ on } \Omega, \text{ then } \int_{\Omega} f dV \leq \int_{\Omega} g dV$$

7.2 Iterated Integrals and Fubini's Theorem

Fubini's Theorem

Fubini's Theorem allows us to connect this abstract definition of integration to our familiar definition of integration in single-variable calculus.

In 2 dimensions, it states that if f is integrable on $R = [a, b] \times [c, d]$, and the sub-integral $\int_c^d f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) dy$ is integrable, then $\int_R f dV = \int_a^b \int_c^d f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) dy dx$.

Furthermore, $\int_a^b \int_c^d f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) dy dx = \int_c^d \int_a^b f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) dx dy$

In n dimensions, this generalizes similarly. If f is integrable on $[a_1, b_1] \times \dots \times [a_n, b_n]$, and all the subintegrals are integrable, then

$$\int_R f(\mathbf{x}) dV = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(\mathbf{x}) dx_n \dots dx_1$$

and the order of the integrals does not matter.

7.3 Polar, Cylindrical, and Spherical Coordinates

There are alternative coordinate systems which can help us solve problems and allow us to exploit symmetry. These are three very specific yet very useful coordinate systems, which we generalize later with the change of variables theorem.

Polar Coordinates

Define $g : [0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2$ by $g\left(\begin{smallmatrix} r \\ \theta \end{smallmatrix}\right) = \begin{bmatrix} r \cos \theta \\ r \sin \theta \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$.

Then, $\int_{g(\Omega)} f\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) dV_{x,y} = \int_{\Omega} f\left(\begin{smallmatrix} r \\ \theta \end{smallmatrix}\right) r dV_{r,\theta}$

It is useful to know that $x^2 + y^2 = r^2$

Visually, we think of r being the distance projected radially outward at an angle of θ which rotates clockwise around the origin.

Cylindrical Coordinates

Define $g : [0, \infty) \times [0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3$ by $g \begin{pmatrix} r \\ \theta \\ z \end{pmatrix} = \begin{bmatrix} r \cos \theta \\ r \sin \theta \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

$$\text{Then, } \int_{g(\Omega)} f \begin{pmatrix} x \\ y \\ z \end{pmatrix} dV_{x,y,z} = \int_{\Omega} f \begin{pmatrix} r \\ \theta \\ z \end{pmatrix} r dV_{r,\theta,z}$$

It is useful to know that $x^2 + y^2 = r^2$

Visually, we think of r being the distance projected radially outward at an angle of θ which rotates clockwise around the origin, at a height of z .

Spherical Coordinates

Define $g : [0, \infty) \times [0, \pi) \times [0, 2\pi) \rightarrow \mathbb{R}^3$ by $g \begin{pmatrix} \rho \\ \phi \\ \theta \end{pmatrix} = \begin{bmatrix} \rho \sin \phi \cos \theta \\ \rho \sin \phi \sin \theta \\ \rho \cos \phi \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

$$\text{Then, } \int_{g(\Omega)} f \begin{pmatrix} x \\ y \\ z \end{pmatrix} dV_{x,y,z} = \int_{\Omega} f \begin{pmatrix} \rho \\ \phi \\ \theta \end{pmatrix} \rho^2 \sin \phi dV_{\rho,\phi,\theta}$$

It is useful to know that $x^2 + y^2 + z^2 = \rho^2$

Visually, we think of ρ being the distance projected radially outward at two angles— ϕ , which starts at the top of the z -axis and rotates down to the bottom, and θ , which rotates all the way about the z -axis.

7.4 Physical Applications

Average value of f on Ω is defined to be $\bar{f} = \frac{1}{\text{vol}(\Omega)} \int_{\Omega} f \, d\text{Vol}$

We can define the integral of a vector *valued* function \mathbf{f} by

$$\int_{\Omega} \mathbf{f} \, d\text{Vol} = \begin{bmatrix} \int_{\Omega} f_1 \, d\text{Vol} \\ \vdots \\ \int_{\Omega} f_n \, d\text{Vol} \end{bmatrix}$$

Average value is still $\bar{\mathbf{f}} = \frac{1}{\text{vol}(\Omega)} \int_{\Omega} \mathbf{f} \, d\text{Vol}$

And *center of mass* of Ω with mass density function $\delta(\mathbf{x})$ is defined to be the

$$\text{point } \bar{\mathbf{x}} = \frac{1}{\text{vol}(\Omega)} \int_{\Omega} \delta(\mathbf{x}) \mathbf{x} \, d\text{Vol}$$

For a rigid body Ω , its *moment of inertia* about an axis l is defined to be

$$\int_{\Omega} \delta(\mathbf{x}) r^2 \, d\text{Vol}, \text{ where } r \text{ is the distance from the point to the axis } l.$$

The gravitational force exerted on the origin by a mass Ω is

$$\mathbf{F} = G \int_{\Omega} \delta(\mathbf{x}) \frac{\mathbf{x}}{\|\mathbf{x}\|^3} \, d\text{Vol}$$

7.5 Determinants and n-Dimensional Volume

The Determinant

For every $n \geq 1$, there is exactly one function $\mathcal{D} : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ times}} \rightarrow \mathbb{R}$ that is

alternating, multilinear, and has $\mathcal{D}(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$

Alternating:

$$\mathcal{D}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) = -\mathcal{D}(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n)$$

Multilinear (linear in all inputs):

$$\mathcal{D}(\mathbf{v}_1, \dots, c\mathbf{v}_i, \dots, \mathbf{v}_n) = c\mathcal{D}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n)$$

and

$$\mathcal{D}(\mathbf{v}_1, \dots, \mathbf{v}_i + \mathbf{v}'_i, \dots, \mathbf{v}_n) = \mathcal{D}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) + \mathcal{D}(\mathbf{v}_1, \dots, \mathbf{v}'_i, \dots, \mathbf{v}_n)$$

Standard:

$$\mathcal{D}(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$$

This function is the determinant. For an $n \times n$ matrix A ,
 $\det A = \mathcal{D}(\mathbf{a}_1, \dots, \mathbf{a}_n)$

Properties

1. If two columns of a matrix are equal, $\det A = 0$
2. Performing row or column operations on a matrix changes the determinant nicely
 1. Swapping two rows or columns leaves $\det A' = -\det A$
 2. Scaling one row or column by c leaves $\det A' = c \det A$
 3. Adding a multiple of one row or column to another leaves $\det A' = \det A$
3. A is nonsingular $\iff \det A \neq 0$
4. $\det(AB) = \det(A) \det(B)$
5. $\det A^{-1} = \frac{1}{\det A}$
6. $\det A^\top = \det A$
7. if A is triangular, $\det A$ is the product of the diagonal entries
8. if T is a linear map, and Ω is a region, $\text{vol}(T(\Omega)) = |\det T| \text{vol}(\Omega)$

Cofactors:

The cofactor $c_{i,j} = (-1)^{i+j} \det A_{i,j}$, where $A_{i,j}$ = the matrix formed by removing row i and column j .

The cofactor matrix C_A is the matrix where $(C_A)_{i,j} = c_{i,j}$

Cramers Rule:

The i^{th} coordinate of \mathbf{x} which solves $A\mathbf{x} = \mathbf{b}$ is given by $x_i = \frac{\det B_i}{\det A}$, where B_i is the matrix obtained by replacing the i th column of A by \mathbf{b}

Inverse matrix "formula":

$$A^{-1} = \frac{1}{\det A} C^\top$$

Permutations:

A permutation $\sigma \in S_n$ is a bijective mapping $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$

Set of all permutations is denoted S_n

The sign of a permutation $\text{sign}(\sigma) = \begin{cases} +1 & \text{odd \# of exchanges} \\ -1 & \text{even \# of exchanges} \end{cases}$

where # of exchanges is the number of swaps required to change the ordered list $(1, \dots, n)$ to $(\sigma(1), \dots, \sigma(n))$

Formulas for the Determinant

Expansion by cofactors:

For any fixed row i

$$\det A = \sum_{j=1}^n a_{i,j} c_{i,j}$$

For any fixed column j

$$\det A = \sum_{i=1}^n a_{i,j} c_{i,j}$$

Sum of row/column permutations:

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}$$

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

7.6 Change of Variables Theorem

Let $\Omega \subset \mathbb{R}^n$ be a region and let there be an open U with $\Omega \subset U$ so that $\mathbf{g} : U \rightarrow \mathbb{R}^n$ is injective, \mathcal{C}^1 , and has a derivative that is mostly invertible (only fails to be invertible on a set of volume 0).

if $f : \mathbf{g}(\Omega) \rightarrow \mathbb{R}$ and $(f \circ \mathbf{g})|\det D\mathbf{g}| : \Omega \rightarrow \mathbb{R}$ are both integrable, then

$$\int_{\mathbf{g}(\Omega)} f(\mathbf{y}) \, d\text{Vol}_{\mathbf{y}} = \int_{\Omega} (f \circ \mathbf{g})(\mathbf{x}) |\det D\mathbf{g}(\mathbf{x})| \, d\text{Vol}_{\mathbf{x}}$$

Chapter 8: Differential Forms and Integration on Manifolds

8.2 Differential Forms

Linear forms

The basis 1-forms $\{dx_1, \dots, dx_n\}$ are defined by $dx_i(v) = v_i$.

The set of linear maps from \mathbb{R}^n to \mathbb{R} is an n dimensional vector space $(\mathbb{R}^n)^*$ spanned by the basis 1-forms $\{dx_1, \dots, dx_n\}$. This space is the *dual space* of \mathbb{R}^n .

Let $I = (i_1, \dots, i_k)$ be a multi-index. Using this, we want to define an alternating, multilinear function $d\mathbf{x}_I : \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{k \text{ times}} \rightarrow \mathbb{R}$ that takes k vectors

in \mathbb{R}^n to \mathbb{R} by $d\mathbf{x}_I(v_1, \dots, v_k) = \det V_I$, where V_I is formed by taking the i_1, \dots, i_k th rows of the matrix $\begin{bmatrix} \uparrow & & \uparrow \\ \mathbf{v}_1 & \cdots & \mathbf{v}_k \\ \downarrow & & \downarrow \end{bmatrix}$. This calculates the volume of

the parallelepiped of the vectors v_1, \dots, v_k projected onto the x_{i_1}, \dots, x_{i_k} plane.

The set of alternating, multilinear functions from $(\mathbb{R}^n)^k$ to \mathbb{R} is denoted by $\Lambda^k(\mathbb{R}^n)^*$. This is the dual space of the exterior power of \mathbb{R}^n . This space is spanned by the basis k-forms which are given by all $d\mathbf{x}_I$ where the (i_1, \dots, i_k) are increasing. The dimension of this space $\dim(\Lambda^k(\mathbb{R}^n)^*) = \binom{n}{k}$. So every $T \in \Lambda^k(\mathbb{R}^n)^*$ is described by $\binom{n}{k}$ coefficients a_I so that

$$T = \sum_{I \text{ increasing}} a_I d\mathbf{x}_I.$$

The wedge product $\wedge : \Lambda^k(\mathbb{R}^n)^* \times \Lambda^l(\mathbb{R}^n)^* \rightarrow \Lambda^{k+l}(\mathbb{R}^n)^*$ can be defined in multiple ways for two forms $\omega = \sum_{I \text{ increasing}} a_I d\mathbf{x}_I \in \Lambda^k(\mathbb{R}^n)^*$ and

$$\eta = \sum_{J \text{ increasing}} b_J d\mathbf{x}_J \in \Lambda^l(\mathbb{R}^n)^*.$$

1.) Concatenation:

For the basis forms, we simply concatenate the multi-indices so that

$d\mathbf{x}_I \wedge d\mathbf{x}_J = d\mathbf{x}_{(I,J)} = d\mathbf{x}_{(i_1, \dots, i_k, j_1, \dots, j_l)}$. By linearity then, we have

$$\omega \wedge \eta = \sum_{I \text{ increasing}} \sum_{J \text{ increasing}} (a_I b_J) d\mathbf{x}_{(I,J)}.$$

2.) Alternization of Tensor Product:

$$\omega \wedge \eta = \text{alt}(\omega \otimes \eta),$$

where \otimes is the tensor product which is defined by

$$(\omega \otimes \eta)(\mathbf{v}_1, \dots, \mathbf{v}_{k+l}) = \omega(\mathbf{v}_1, \dots, \mathbf{v}_k) \cdot \eta(\mathbf{v}_{k+1}, \dots, \mathbf{v}_{k+l})$$

and alt is the alternization for a multilinear $\gamma : (\mathbb{R}^n)^k \rightarrow \mathbb{R}$ defined by

$$\text{alt}(\gamma)(\mathbf{v}_1, \dots, \mathbf{v}_k) = \sum_{\sigma \in S_k} \text{sign}(\sigma) \gamma(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(k)}).$$

Note that every $d\mathbf{x}_I = d\mathbf{x}_{i_1} \wedge \dots \wedge d\mathbf{x}_{i_k}$

Properties:

$$\omega \in \Lambda^k(\mathbb{R}^n)^*, \eta \in \Lambda^l(\mathbb{R}^n)^*, \mu \in \Lambda^m(\mathbb{R}^n)^*$$

1. Bilinear

$$1. (\omega + \eta) \wedge \mu = \omega \wedge \mu + \eta \wedge \mu$$

$$2. (c\omega) \wedge \eta = c(\omega \wedge \eta)$$

2. Skew-commutative

$$1. \omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$$

3. Associative

$$1. (\omega \wedge \eta) \wedge \mu = \omega \wedge (\eta \wedge \mu)$$

Differential Forms

A differential form associates a function to each of the $\binom{n}{k}$ basis forms. A

differential k -form on \mathbb{R}^n is an expression $\omega = \sum_{I \text{ increasing}} f_I(\mathbf{x}) d\mathbf{x}_I$.

Special cases:

A 0-form is a smooth function.

An n -form is an expression of the form $\omega = f(\mathbf{x}) d\mathbf{x}_1 \wedge \dots \wedge d\mathbf{x}_n$.

The set of k forms on \mathbb{R}^n forms a vector space $\mathcal{A}^k(\mathbb{R}^n)$

Properties:

$\omega \in \mathcal{A}^k(\mathbb{R}^n), \eta \in \mathcal{A}^l(\mathbb{R}^n), \mu \in \mathcal{A}^m(\mathbb{R}^n)$

1. Distributes (when $k = l$)

$$1. (\omega + \eta) \wedge \mu = \omega \wedge \mu + \eta \wedge \mu$$

2. Skew-commutative

$$1. \omega \wedge \eta = (-1)^{kl} \eta \wedge \omega$$

3. Associative

$$1. (\omega \wedge \eta) \wedge \mu = \omega \wedge (\eta \wedge \mu)$$

4. 1. Associative / commutative (when $k = l = m$)

$$2. (\omega + \eta) + \mu = \omega + (\eta + \mu)$$

$$3. \omega + \eta = \eta + \omega$$

Exterior derivative

The exterior derivative d allows us to differentiate differential forms. The exterior derivative turns k forms into $k + 1$ forms.

For 0-forms, $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ we want $df(\mathbf{x}) = Df(\mathbf{x})$ as a linear map $\mathbb{R}^n \rightarrow \mathbb{R}$. This motivates us to define

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

For a general k form $\omega = \sum_I f_I(\mathbf{x}) d\mathbf{x}_I$, we define

$$d\omega = \sum_I df_I \wedge d\mathbf{x}_I = \sum_I \sum_{j=1}^n \frac{\partial f_I}{\partial x_j} dx_j \wedge d\mathbf{x}_{i_1} \wedge \cdots \wedge d\mathbf{x}_{i_k}$$

Properties:

$\omega \in \mathcal{A}^k(\mathbb{R}^n), \eta \in \mathcal{A}^l(\mathbb{R}^n), f \in \mathcal{A}^0(\mathbb{R}^n)$ (smooth function)

1. Distributes (when $k = l$)

1. $d(\omega + \eta) = d\omega + d\eta$
2. Scalar function product rule
 1. $d(f\omega) = df \wedge \omega + f d\omega$
3. General product rule
 1. $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta$
4. Derivative twice is 0
 1. $d(d\omega) = 0$

Pullback

Pullback is a sort of differential-forms generalization of the chain rule.

If (for open U) $g : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ is smooth, and $\omega \in \mathcal{A}^k(\mathbb{R}^n)$, then we define $g^*\omega \in \mathcal{A}^k(U)$ in multiple steps

To pull back a 0-form (function), we compose by

$$g^*f = f \circ g$$

To pull back a basis 1-form, we set

$$g^*dx_i = dg_i = \sum_{j=1}^m \frac{\partial g_i}{\partial u_j} du_j$$

and to pull back a basis k-form, we set

$$g^*(dx_{i_1} \wedge \cdots \wedge dx_{i_k}) = g^*dx_{i_1} \wedge \cdots \wedge g^*dx_{i_k} = dg_{i_1} \wedge \cdots \wedge dg_{i_k} = dg_I$$

and finally, to take the pullback of a general form, we distribute the pullback so,

$$g^*\omega = g^*\left(\sum_I f_I(\mathbf{x}) d\mathbf{x}_I\right) = \sum_I (f_I \circ g)(\mathbf{x}) dg_I = \sum_I (f_I \circ g)(\mathbf{x}) dg_{i_1} \wedge \cdots \wedge dg_{i_k}$$

This tells us that

$$\mathbf{g}^* d\mathbf{x}_I = \sum_{J \text{ increasing}} \det\left[\frac{\partial \mathbf{g}_I}{\partial \mathbf{u}_J}\right] d\mathbf{u}_J = \sum_{J \text{ increasing}} \begin{vmatrix} \frac{\partial g_{i_1}}{\partial u_{j_1}} & \cdots & \frac{\partial g_{i_1}}{\partial u_{j_k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{i_k}}{\partial u_{j_1}} & \cdots & \frac{\partial g_{i_k}}{\partial u_{j_k}} \end{vmatrix} du_{j_1} \wedge \cdots \wedge du_{j_k}$$

also note that $\mathbf{g}^*(d\omega) = d(\mathbf{g}^*\omega)$

Defining integration

Given an n -form $\omega = f(\mathbf{x}) = dx_1 \wedge \cdots \wedge dx_n$ on a region $\Omega \subset \mathbb{R}^n$, we define

$$\int_{\Omega} \omega := \int_{\Omega} f \, d\text{Vol}$$

With this definition, we can nicely restate the change of variables theorem in a nice way. If $\mathbf{g} : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is smooth, injective, and has $\det(D\mathbf{g}) > 0$ on a region Ω , then for any $\omega \in \mathcal{A}^n(\mathbb{R}^n)$ we have $\int_{\mathbf{g}(\Omega)} \omega = \int_{\Omega} \mathbf{g}^*\omega$. This motivates us to define integration on a manifold.

If $\mathbf{g} : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is smooth, injective, and has $\text{rank}(\det D\mathbf{g}) = k$ on a region Ω , and only doesn't have this rank on a set of volume 0, then for a parameterized k -dimensional manifold $M = \mathbf{g}(\Omega) \subset \mathbb{R}^n$ and for any $\omega \in \mathcal{A}^k(\mathbb{R}^n)$ we define

$$\int_M \omega := \int_{\Omega} \mathbf{g}^*\omega$$

This is well defined, and has the same result for different parameterizations (as long as both parameterizations orient M in the same way)

8.3 Line Integrals and Green's Theorem

We can associate a 1-form in \mathbb{R}^n ($\omega = \sum f_i dx_i$) with a vector field

$\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, where $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix}$. Then, if we have a curve \mathcal{C} ,

parameterized by some function $\mathbf{g} : [a, b] \rightarrow \mathcal{C}$, we can find out how much *work* is done on a particle traveling on the path \mathcal{C} by the force field \mathbf{F} .

This work is given by the line integral $\int_{\mathcal{C}} \omega = \int_{\mathcal{C}} f_1(\mathbf{x}) dx_1 + \dots + f_n(\mathbf{x}) dx_n$.

To compute this integral, we pull back via the parameterization \mathbf{g} to get

$$\int_{\mathbf{g}([a,b])} \omega = \int_{[a,b]} \mathbf{g}^*(\omega) = \int_a^b f_1(\mathbf{g}(t)) g'_1(t) dt + \dots + f_n(\mathbf{g}(t)) g'_n(t) dt = \int_a^b \mathbf{F}(\mathbf{g}(t)) \cdot \mathbf{g}'(t) dt$$

It's common notation to let $\mathbf{s} = \mathbf{g}(t)$ so that

$$\int_a^b \mathbf{F}(\mathbf{g}(t)) \cdot \mathbf{g}'(t) dt = \int_a^b \mathbf{F}(\mathbf{s}) \cdot d\mathbf{s}$$

F.T.C for line integrals

For a 1-form ω , if we can find a potential 0-form f so that $\omega = df$, then for any curve \mathcal{C} with endpoints $\mathbf{A}, \mathbf{B} \in \mathcal{C}$, we have $\int_{\mathcal{C}} \omega = f(\mathbf{A}) - f(\mathbf{B})$. In

traditional vector notation, if we can find an f so that the force field $\mathbf{F} = \nabla f$, then for a curve \mathcal{C} parameterized by a $\mathbf{g} : [a, b] \rightarrow \mathbb{R}^n$, we have

$$\int_{\mathcal{C}} \mathbf{F}(\mathbf{s}) \cdot d\mathbf{s} = f(\mathbf{g}(b)) - f(\mathbf{g}(a)).$$

Exactness

A 1-form ω is exact if any of the following equivalent statements are true:
(The endpoints of a curve \mathcal{C} are denoted \mathbf{A} and \mathbf{B})

1. $\omega = df$ for some 0-form f
2. For every closed curve (when $\mathbf{A} = \mathbf{B}$), we have $\int_{\mathcal{C}} \omega = 0$

3. $\int_C \omega$ is path-independent. ($\int_C \omega = \int_{C'} \omega$ if C and C' have the same endpoints)

Closedness

A 1-form ω is closed if $d\omega = 0$

Simply Connected Regions

- exactness \implies closedness
 - because $d^2 = 0$
- closedness $\not\Rightarrow$ exactness
 - *except on special regions called "simply connected regions"
- $X \subset \mathbb{R}^n$ is called *simply connected* if any pair of points can be joined by a piecewise \mathcal{C}^1 path (connected), and every closed curve which doesn't intersect itself (every simple curve) in X can be continuously shrunk to a point in X .
 - *Intuition*: the subset has no holes
 - Think about an open set. You can draw any closed curve and shrink it. Think about an open set with a hole. If you draw a curve around the hole, you can't shrink the curve to a point anymore
- If ω is a smooth, closed 1-form on a simply connected region $\Omega \subset \mathbb{R}^2$, then ω is exact.

Green's Theorem

Green's Theorem for a Rectangle:

For a rectangle $R \subset \mathbb{R}^2$, and a 1-form ω on R .

$$\int_{\partial R} \omega = \int_R d\omega$$

(Where ∂R is counterclockwise)

Classical Green's Theorem:

For an $\mathbf{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, where $\mathbf{F}(x, y) = \begin{bmatrix} F_1(x, y) \\ F_2(x, y) \end{bmatrix}$. And for a region D bounded by C . Then

$$\int_C F_1(x, y) dx + F_2(x, y) dy = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy$$

Slight Generalization:

- If $S \subset \mathbb{R}^2$ is parameterized by a rectangle, then $\int_{\partial S} \omega = \int_S d\omega$.
- Green's Theorem holds for any region S that can be decomposed as a finite union of parametrized rectangles which overlap only along their boundaries.

Winding Number

Something fun we can do is count the number of times a parameterized curve γ "winds" around the origin. We define

$\text{wind}(\gamma) = \frac{1}{2\pi} \int_{\gamma} \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy$, and this counts the number of times a curve loops counterclockwise around the origin.

8.4 Surface Integrals and Flux

Parameterized Surface

We can work with special 2-dimensional surfaces in \mathbb{R}^3 , like spheres, toruses, cylinders, etc.

Suppose we have a surface $S = \mathbf{g}(U)$ parameterized by an injective function $\mathbf{g} : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^n$ (for a bounded, open U) that always has $\text{rank}(D\mathbf{g}(\mathbf{x})) = 2$. Then S is a *parameterized surface*.

We define the integral of a 2-form $\omega \in \mathcal{A}^2(\mathbb{R}^n)$ over this parameterized surface to be $\int_S \omega = \int_U g^* \omega$.

Orientation

For each tangent plane of S , find a positively oriented basis (\mathbf{u}, \mathbf{v}) (i.e. one where $\det(\mathbf{u}, \mathbf{v}) > 0$). An orientation for S is a continuously varying notion of what a positively oriented basis for the plane at each point of S should be.

An orientation can be determined in multiple ways.

S has an orientation iff we can choose various parameterizations g_i of subsets $S_i \subset S$ such that $\bigcup_i S_i = S$ and that at every point $\mathbf{x} \in S$ the corresponding g_i has $\det\left(\frac{\partial g_i}{\partial u_1}, \frac{\partial g_i}{\partial u_2}\right) > 0$ (the tangent vectors of the parameterization form a positively oriented basis).

An orientation can also be determined if we have a 2-form ω which is never 0 on S . This is because at each point $\mathbf{x} \in S$, we can pick tangent vectors \mathbf{u}, \mathbf{v} so that $\omega(\mathbf{x})(\mathbf{u}, \mathbf{v}) > 0$, which defines a continuously varying notion of what a positively oriented basis for the plane should be.

Area Forms

For an oriented surface S , its oriented area 2-form σ assigns at each point $\mathbf{x} \in S$ the signed area of the parallelogram that the tangent vectors at that point span.

In \mathbb{R}^3 (where $\mathbf{x} = (x, y, z)$), the area 2-form gives the *surface area* of a surface S . If at each point $\mathbf{x} \in S$, the surface S has outward pointing unit

normal given by $\mathbf{n}(\mathbf{x}) = \begin{bmatrix} n_1(\mathbf{x}) \\ n_2(\mathbf{x}) \\ n_3(\mathbf{x}) \end{bmatrix}$, then the corresponding area 2-form is

$\sigma = n_1(\mathbf{x}) dy \wedge dz + n_2(\mathbf{x}) dz \wedge dx + n_3(\mathbf{x}) dx \wedge dy$. This works because \mathbf{n} is

a unit vector, so the area of the plane spanned by tangent vectors \mathbf{u} and \mathbf{v} is

given by $\det \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \mathbf{n} & \mathbf{u} & \mathbf{v} \\ \downarrow & \downarrow & \downarrow \end{pmatrix}$.

Flux

// TODO

8.5 Stoke's Theorem

// TODO

8.6 Applications to Physics

8.7 Applications to Topology

Chapter 9: Eigenvalues, Eigenvectors, and Applications

9.1 Linear Transformations and Change of Basis

// TODO

9.2 Eigenvalues, Eigenvectors, and Diagonalizability

// TODO

9.3 Difference Equations and Ordinary Differential Equations

// TODO

9.4 The Spectral Theorem

// TODO