Notes for 3510

This is a big reference sheet 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, as well as some fun results.

Chapter 6: Solving Nonlinear Problems

6.1 The Contraction Mapping Principle

Convergence of Vector Sums

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

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

Contraction Mappings

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

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

$$\|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{y})\| \leq c\|\boldsymbol{x} - \boldsymbol{y}\|$$

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 princple. The contraction mapping principle states that if you call a contraction mapping $\boldsymbol{f}$ enough times, then it eventually converges on some point, called the fixed point of $\boldsymbol{f}$.

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

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 $\dfrac{f(b) - f(a)}{b - a} = f'(x)$ 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 is how we generalize the MVT to $\mathbb{R}^n$

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

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

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

Matrix Theorems

Generalization of a geometric series. If $A \in \mathcal{M}_{n \times n}$ and $\|A\|_{\mathrm{op}} < 1$, then $\sum_{i=0}^{\infty} A^i = (I - A)^{-1}$

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

Since $\mathcal{M}_{m \times n}$ is isomorphic to $\mathbb{R}^{nm}$, we can take derivative maps of matrix-valued functions. For example, consider $f : \mathcal{M}_{n \times n} \to \mathcal{M}_{n \times n}$ 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 $[Df(A)] : \mathbb{R}^{n^2} \to \mathbb{R}^{n^2}$, 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 $\mathrm{O}(n) \subset \mathcal{M}_{n \times n} = \{\text{set of matrices where } A^T A = I\}$ is a $\frac{1}{2}n(n-1)$-dimensional submanifold, and $T_I \mathrm{O}(n)$ (tangent plane at the identity) is the set of skew-symmetric matrices $\mathrm{Skew}(n) = \{\text{set of matrices where } A^T = -A\}$

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 $\boldsymbol{f}$ has a local inverse, and what the derivative of that inverse looks like. To have an inverse, the derivative map must be locally invertable, 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, $\boldsymbol{f} : U \to \mathbb{R}^n$ is $\mathscr{C}^1$, $\boldsymbol{x}_0 \in U$ and $D\boldsymbol{f}(\boldsymbol{x}_0)$ is invertible, then $\exists$ two open neighborhoods, $V \subset U$ of $\boldsymbol{x}_0$ and W of $\boldsymbol{y}_0 := \boldsymbol{f}(\boldsymbol{x}_0)$, and a $\mathscr{C}^1$ function between them $\boldsymbol{g} : V \to W$ so that

$\forall \boldsymbol{x} \in V, \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) = \boldsymbol{x}$
$\forall \boldsymbol{y} \in W, \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{y})) = \boldsymbol{y}$

Furthermore, $D\boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x})) = (D\boldsymbol{f}(\boldsymbol{x}))^{-1}$

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 $\mathscr{C}^1$ function $\boldsymbol{F} : U \subset \mathbb{R}^n \to \mathbb{R}^m$ where U is open. Then, writing a vector $\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} \in \mathbb{R}^n$ with $\boldsymbol{x} \in \mathbb{R}^{n-m}$ and $\boldsymbol{y} \in \mathbb{R}^m$, suppose $\boldsymbol{F} \begin{pmatrix} \boldsymbol{x}_0 \\ \boldsymbol{y}_0 \end{pmatrix} = \vec{\boldsymbol{0}}$ describes a series of m implicit equations and $\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{y}} \begin{pmatrix} \boldsymbol{x}_0 \\ \boldsymbol{y}_0 \end{pmatrix} \in \mathcal{M}_{m \times m}$ is invertible.

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

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

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

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

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: paramaterized by an injective funtion that never has velocity 0

more formally

1. For any $\boldsymbol{p} \in M$, $\exists$ a neighborhood $W \subset \mathbb{R}^n$ of $\boldsymbol{p}$ so that $W \cap M$ is the graph of a smooth function $\boldsymbol{f} : V \subset \mathbb{R}^k \to \mathbb{R}^{n-k}$, where V is open. Remember that

$\text{graph}\{\boldsymbol{f}(\boldsymbol{x})\} = \{\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{f}(\boldsymbol{x}) \end{bmatrix} : \boldsymbol{x} \in \mathbb{R}^k\}$, but that the order of the x_i's doesn't matter.

2. For any $\boldsymbol{p} \in M$, $\exists$ a neighborhood $W \subset \mathbb{R}^n$ of $\boldsymbol{p}$ and a smooth function $\boldsymbol{F} : W \to \mathbb{R}^{n-k}$ so that the preimage (level set) $\boldsymbol{F}^{-1}(\vec{\boldsymbol{0}}) = W \cap M$ and $\text{rank}\{D\boldsymbol{F}(\boldsymbol{x})\} = n - k$ for every $\boldsymbol{x} \in W \cap M$.

3. For any $\boldsymbol{p} \in M$, $\exists$ a neighborhood $W \subset \mathbb{R}^n$ of $\boldsymbol{p}$ so that $W \cap M$ is the image of a smooth function $\boldsymbol{g} : U \subset \mathbb{R}^k \to \mathbb{R}^n$, where U is open. $\boldsymbol{g}$ must be injective, have $\text{rank}\{D\boldsymbol{g}(\boldsymbol{u})\} = k$ for all $\boldsymbol{u} \in U$, and have $\boldsymbol{g}^{-1} : W \cap M \to U$ be continuous.

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

1. If M is locally at $\boldsymbol{p}$ a graph of $\boldsymbol{f}$ with $\boldsymbol{p} = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{f}(\boldsymbol{a}) \end{bmatrix}$, then $T_{\boldsymbol{p}}M = \text{graph}\{D\boldsymbol{f}(\boldsymbol{a})\}$.

2. If M is locally a preimage of $\boldsymbol{F}$, then $T_{\boldsymbol{p}}M = \text{ker}\{D\boldsymbol{f}(\boldsymbol{p})\}$.

3. If M is locally paramaterized by $\boldsymbol{g}$ with $\boldsymbol{p} = \boldsymbol{g}(\boldsymbol{a})$, then $T_{\boldsymbol{p}}M = \text{im}\{D\boldsymbol{g}(\boldsymbol{a})\}$

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 $\mathcal{P} := \mathcal{P}_1 \times \ldots \times \mathcal{P}_n$ (each $\mathcal{P}_i$ for $1 \le i \le n$ 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-jth rectangle of R. Although for $\mathbb{R}^n$ it should be more like the $(i, j, k, \ldots, n)$-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 $f : R \to \mathbb{R}$, $\sup_{\boldsymbol{x} \in S} f(\boldsymbol{x})$ is the "least upper bound" and $\inf_{\boldsymbol{x} \in S} f(\boldsymbol{x})$ 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 $m_{i,j} := \inf_{\boldsymbol{x} \in R_{i,j}} f(\boldsymbol{x})$ and $M_{i,j} := \sup_{\boldsymbol{x} \in R_{i,j}} f(\boldsymbol{x})$, 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 $\displaystyle\int_R f \, dV := I$.

Properties of Integrals