140C Summary of Definitions and Results

1 Linear Transformations

1.1 Some facts about vector spaces

Definition 1.1 *The span of a subset X of $\mathbb{R}^n$ is the set, say E, of all linear combinations of X, we say X spans E.*

Definition 1.2 *The dimension of a vector space X, denoted $dim(X)$, is the smallest integer r so that any $r+1$ vectors are linearly dependent.*

Definition 1.3 *A basis for a vector space X is an independent collection of vectors that spans X.*

Theorem 1.4 *If the vector space X is spanned by r vectors, then $dim(X) \leq r$.*

Corollary 1.5 *$dim(\mathbb{R}^n) = n$.*

Theorem 1.6 *If the vector space X has $dim(X) = n$, then a set of n vectors spans X if and only if it is independent.*

Corollary 1.7 *If X is a vector space with $dim(X) = n$, then X has a basis and every basis has n vectors.*

1.2 Linear transformations

Definition 1.8 *A mapping $A : X \to Y$ is a linear transformation if $\forall x_1, x_2 \in X$, all scalars c we have $A(x_1 + x_2) = Ax_1 + Ax_2$ and $A(cx_1) = cAx_1$.*

Definition 1.9 *The set of all linear transformations from X to Y is denoted $L(X, Y)$.*

Definition 1.10 *If $A \in L(X, X)$, we may call it a linear operator.*

Remark 1.11 *If A is a linear operator on X, one-to-one, onto, we say that A is invertible. There exists $A^{-1} : X \to X$ with $A^{-1}(Ax) = x$. Moreover, $A(A^{-1}x = x)$ and $A^{-1} \in L(X, X)$*

Theorem 1.12 *If A is a linear operator on a finite dimensional vector space X, it is one-to-one if and only if its range is all of X.*

Definition 1.13 *If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$, the norm of A is $\|A\| := \sup_{|x|\leq 1} |Ax|$. We have $|Ax| \leq \|A\||x|$. Moreoever, if $|Ax| \leq \lambda|x|$ then $\|A\| \leq \lambda$.*

Theorem 1.14 *If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ then $\|Ax\| < \infty$ and A is uniformly continuous.*

Fact 1.15 *$\|A + B\| \leq \|A\| + \|B\|$ and $\|cA\| = |c|\|A\|$.*

Fact 1.16 *$\|AB\| \leq \|A\|\|B\|$.*

Theorem 1.17 *If $\|B - A\|\|A^{-1}\| < 1$, then B is invertible.*

Theorem 1.18 *If A is invertible then the map $A \mapsto A^{-1}$ is continuous.*

Fact 1.19 *Let S be a metric space and $\left(a_{i,j}\right)_{i=1,j=1}^{i=m,j=n}$ be real and continuous* functions *on S. Suppose that for all $p \in S$, $A_p \in L(\mathbb{R}^n, \mathbb{R}^m)$ has matrix entries $a_{i,j}(p)$. Then the mapping $p \mapsto A_p$ is a continuous mapping of S into $L(\mathbb{R}^n, R^m)$.*

2 Differentiation

2.1 Basics

Definition 2.1 *Consider $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$, where E is open and let $x \in E$. If there exists $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ with*

$$\lim_{h \to 0} \frac{|f(x+h) - f(x) - Ah|}{|h|} = 0$$

then f is differentiable at x and we write $f'(x) = A$. If f is differentiable at every $x \in E$, we say it is differentiable in E.

Fact 2.2 *If a function f is differentiable at x, then its derivative $f'(x)$ is unique.*

Theorem 2.3 (Chain Rule) *Let $E \subset \mathbb{R}^n$ be open and suppose $f : E \to \mathbb{R}^m$ is differentiable at x_0 in E. Suppose further that $g : \mathbb{R}^m to \mathbb{R}^k$ is differentiable at $f(x_0)$. Then, the map $F : E \to \mathbb{R}^k$ given by $F(x) = g(f(x))$ is differentiable at x_0 and*

$$F'(x_0) = g'(f(x_0))f'(x_0).$$

2.2 Partial derivatives

Definition 2.4 (Partial Derivatives) $(D_j f_i)(x) := \lim_{t \to 0} \frac{f_i(x+te_j) - f_i x}{t}$, *provided the limit exists. One often writes $\frac{\partial f_i}{\partial x_j}$ in place of $(D_j f_i)$.*

Fact 2.5 (Warning) *The existence of all partial derivatives does not imply differentiability in the sense of Definition 2.1.*

Theorem 2.6 *If $E \subset \mathbb{R}^n$ is open and $f : E \to \mathbb{R}^m$ is differentiable at x then the partial derivatives $D_j f_i$ exist and*

$$f'(x)e_j = \sum_{i=1}^{m} (D_j f_i)(x) u_i$$

Remark 2.7 *Essentially this is saying that the jth column of the matrix $[f'(x)]$ consists of the entries $(D_j f_i)(x)$, $i = 1, ..., n$. In other words, the matrix associated with the linear transformation $f'(x)$ is populated by the partial derivatives evaluated at x.*

Definition 2.8 (Gradient) *Let $f : E \subset \mathbb{R}^n \to \mathbb{R}$ where E is open. Associated with every x in E is a vector called the gradient of f at x given by*

$$(\nabla f)(x) := \sum_{j=1}^{n} (D_j f)(x) e_j.$$

Definition 2.9 (Directional derivative) *Given a function $f : E \subset \mathbb{R}^n \to \mathbb{R}$ (where E is open), and a unit-norm vector $u \in \mathbb{R}^n$, the quantity $\lim_{t\to 0} \frac{f(x+tu)-f(x)}{t}$ (when the limit exists) is called the directional derivative of f at x, in the direction of u. It is denoted $(D_u f)(x)$.*

Fact 2.10 *The directional derivative satisfies $\lim_{t\to 0} \frac{f(x+tu)-f(x)}{t} = (\nabla f)(x) \cdot u$*

Theorem 2.11 *Suppose $E \subset \mathbb{R}^n$ is* convex *and* open *and let $f : E \to \mathbb{R}^m$. Suppose there exists a number M such that $\|f'(x)\| \leq M$ for all $x \in E$. Then*

$$|f(b) - f(a)| \leq M|b - a| \quad \text{for all} \quad a, b \in E.$$

Definition 2.12 *A differentiable mapping f of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^m$ is continuously differentiable in E if $f'(x)$ is a continuous mapping of E into $L(\mathbb{R}^n, \mathbb{R}^m)$.*
In other words, we say that f is continuously differentiable in E if for every $x \in E$ and $\epsilon > 0$, there exists a $\delta > 0$ such that $\|f'(x) - f'(y)\| < \epsilon$ for every $y \in E$ with $|y - x| < \delta$.

In this case, we say that f is C^1 or $f \in C^1(E)$.

Theorem 2.13 *Suppose $f : E \subset \mathbb{R}^n \to \mathbb{R}^m$ (where E is open). Then $f \in C^1(E)$ if and only if the partial derivatives $D_j f_i$ exist and are continuous on E for $i = 1, ..., m$, and $j = 1, ..., n$.*

3 The inverse function theorem

Theorem 3.1 (The contraction principle) *Let X be a complete metric space, with metric d. Suppose that $\phi : X \to X$ with $d(\phi(x), \phi(y)) \leq cd(x, y)$ for some $c < 1$ and all $x, y \in X$. Then there exists a unique $x \in X$ with $x = \phi(x)$.*

Theorem 3.2 *Let f be a C^1 mapping from an open set $E \subset \mathbb{R}^n$ to $\mathbb{R}^n$. Suppose that $f(a) = b$ for some $a \in E$ and that $f'(a)$ is invertible. Then*

(a) There exist open sets $U \in \mathbb{R}^n$ and $V \in \mathbb{R}^n$ with $a \in U$ and $b \in V$ such that the restriction $f : U \to V$ is one-to-one and onto.

(b) Denote by $g : V \to U$, the inverse of f, then $g \in C^1(V)$.

Remark 3.3 *Note the local nature of the inverse function theorem. The function $f \in C^1(E)$ is only one-to-one and onto around a where $f'(a)$ is invertible, and its inverse is also C^1. This does not mean that f needs to be invertible on its entire domain.*

Corollary 3.4 *If $f : E \subset \mathbb{R}^n \to \mathbb{R}^n$, where E is open, is a C^1 map, and if f' is invertible for every $x \in E$ then $f(W)$ is an open subset of $\mathbb{R}^n$ for every open set $W \subset E$.*

4 The implicit function theorem

Definition 4.1 (notation) *Suppose $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^m)$. Then we define $A_x \in L(\mathbb{R}^n, \mathbb{R}^n)$ via its action on vectors $h \in \mathbb{R}^n$: $A_x h = A(h, 0)$. We define $A_y \in L(\mathbb{R}^m, \mathbb{R}^n)$ via its action on vectors $k \in \mathbb{R}^m$: $A_y k = A(0, k)$.*

Theorem 4.2 (The implicit function theorem – the linear case) *Suppose $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ and A_x is invertible. Then to every $k \in \mathbb{R}^m$ there exists a unique $h \in \mathbb{R}^n$ such that A(h,k) = 0. Moreover $h = -A_x^{-1} A_y k$.*

Theorem 4.3 (The implicit function theorem) *Let f be a C^1 mapping from an open set $E \subset \mathbb{R}^{n+m}$ to $\mathbb{R}^n$. Suppose that $f(a, b) = 0$ where $a \in \mathbb{R}^n, b \in \mathbb{R}^m$, and $(a, b) \in E$. Denoting by $A := f'(a, b)$, suppose that A_x is invertible.*

Then, there exist open sets $U \in \mathbb{R}^{n+m}$ and $W \in \mathbb{R}^m$ with $(a, b) \in U$ and $b \in W$ such that

(a) For every $y \in W$ there exists a unique x with $(x, y) \in U$ and $f(x, y) = 0$.

(b) Writing $g(y) := x$, g *is a* C^1 *map of* W *into* $\mathbb{R}^n$ *satisfying* $g(b) = a$, $f(g(y), y) = 0$ *and* $g'(b) = -(A_x)^{-1}A_y$.