Reference Sheet for Final Exam

Erica Shannon
Math 211, Irena Swanson
(Dated: December 12, 2007)

Completeness of Reals Every bounded monotonic sequence $\{x_n\} \in \mathbb{R}$ has a unique limit $a \in \mathbb{R}$. Every nonempty subset $A \subset \mathbb{R}$ that is bounded above has a sup or in $\mathbb{R}$. Every nonempty compact subset of $\mathbb{R}^n$ attains its maximum.

Linear Combinations and Bases A **linear combination** of vectors $v_1, \ldots, v_k$ is any vector of the form $c_1v_1 + \ldots + c_kv_k$ for $c_i \in \mathbb{R}$. A **basis** for a vector space V is a set of finite vectors such that every vector in V can be written *uniquely* as a linear combination of them. The **standard basis** is $e_1, \ldots, e_n$.

Vector Angle Formula

$$\cos\theta_{x,y} = \frac{\langle x, y\rangle}{|x||y|} \in [-1, 1] \text{ for } \theta \in [0, \pi]$$

Vector Shadows For a the vector projection of v onto w,

$$a = \frac{\langle v, w\rangle}{|w|^2} w$$

Sequence Lemma and Bolzano-Weierstrass Theorem Every bounded sequence $\{x_\nu\} \in \mathbb{R}$ has a convergent subsequence. For $A \subset \mathbb{R}$, A is bounded iff every sequence in A has a subsequence which converges in $\mathbb{R}^n$.

Useful Inequalities

- Cauchy-Schwarz: $\langle x, y\rangle \leq \langle x, x\rangle\langle y, y\rangle$
- Triangle Inequality: $||x| - |y|| \leq |x + y| \leq |x| + |y|$
- Size Bounds: $|x_j| \leq |x| \leq \sum_{i=1}^n |x_i|$

Continuity f is continuous at a iff $\{x_\nu\} \to a \Rightarrow \{f(x_\nu)\} \to f(a)$.

Random Topology A is closed if it contains all of its limit points. An arbitrary intersection of closed sets is closed. An arbitrary union of closed sets need not be closed. A finite union of closed sets is closed. The null set and $\mathbb{R}^n$ are both closed and open. A set is open if its complement is closed. An arbitrary union of open sets is open. A finite intersection of open sets is open.

Linear Transformation Matrices Matrix of a linear T, where $T(x) = (\langle x, a_1\rangle, \langle x, a_2\rangle, \ldots)$:

$$T = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} T(e_1) & \cdots & T(e_n) \end{bmatrix}$$

For compositions of linear transformations, multiply matrices.

Linear Transformation "Workhorse" Theorem Let $T : \mathbb{R}^n \to \mathbb{R}$ be linear. Then there exists $a \in \mathbb{R}^n$ such that $\forall x \in \mathbb{R}^n$, $T(x) = \langle x, a\rangle$, and $a_i = T(e_i)$ for all $i = \{1, \ldots, n\}$

Invertibility An $m \times n$ matrix is **invertible** iff there exists an $n \times m$ matrix B such that $A \cdot B = I_m$ and $B \cdot A = I_n$. If an $m \times n$ matrix is invertible, then its reduced echelon form is I_n, and $m = n$. Also, $(MN)^{-1} = N^{-1}M^{-1}$.

Elementary Matrices

- Recombine: $R_{i;j,a}$ makes the new i^{th} row = the old i^{th} row + $a\cdot$ old j^{th} row
- Scale: $S_{i,a}$ makes the new i^{th} row = $a\cdot$ old i^{th} row
- Transpose: $T_{i,j}$ switches i and j rows

Taylor Polynomials

$$T_n(x) = \sum_{i=0}^{n} \frac{f^i(a)(x-a)}{i!}$$

$$R_n(x) = f^{n+1}(c)\frac{(x-a)^{n+1}}{(n+1)!} \text{ for some } c \in (a, x) \cup (x, a)$$

Interior Point Let $A \subseteq \mathbb{R}^n$. Then $a \in A$ is in the **interior** of A iff there exists $\epsilon > 0$ such that $B(a, \epsilon) \subseteq A$.

Differentiability For $f : A(\subseteq \mathbb{R}^n) \to \mathbb{R}^m$, a in interior of A, we say that f is differentiable at a if there exists a linear mapping $T_a : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h\to 0} \frac{|f(a+h) - f(a) - T_a(h)|}{|h|} = 0.$$

T_a is the *unique* derivative of f at a, written Df_a or $(Df)_a$, and has a corresponding Jacobian matrix $f'(a)$ which is an $m \times n$ matrix. This matrix sometimes exists even where f may not be differentiable. Also, f differentiable at $a \Leftrightarrow$ each component function of f is differentiable at a.

Linear Boundedness If $T : A(\subseteq \mathbb{R}^n) \to \mathbb{R}^m$ is linear, then $\exists c \in \mathbb{R} : \forall x \in \mathbb{R}^n, |T(x)| \leq c|x|$.

Product & Quotient Rules If $f, g : A \to \mathbb{R}$, a in interior of A, f, g differentiable at a, then fg is differentiable with

$$D(fg)(a) = f(a)Dg_a + g(a)Df_a$$

If $g(a) \neq 0$, then $\frac{f}{g}$ if differentiable with

$$D(\frac{f}{g})_a = \frac{g(a)Df_a - f(a)Dg_a}{g(a)^2}$$

Chain Rule If $f : A \to \mathbb{R}^m$ is differentiable at a in the interior of A, and $g : f(A) \to \mathbb{R}^l$ differentiable at $f(a)$ in the interior of $f(A)$, then $g \circ f$ is also differentiable at a and

$$D(g \circ f)_a = Dg_{f(a)} \circ Df_a$$

. . . in Coordinates Let $f : a(\subseteq \mathbb{R}^n) \to \mathbb{R}^m$ differentiable at a in the interior of A, and $g : f(A) \to \mathbb{R}$ differentiable at $f(a)$ in the interior of $f(A)$. Then

$$D_j(g \circ f) = \sum_{i=1}^{m} D_i g(f(a)) \cdot D_j f_i(a).$$

In other words,

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$$

Partial & Directional Derivatives Let $f : A(\subseteq \mathbb{R}^n) \to \mathbb{R}^m, a \in A, j = \{1, \dots, n\}$. Then the j^{th} partial derivative of f at a (if it exists) is

$$D_j f_a = \frac{\partial f}{\partial x_j}(a) = \lim_{t \to 0, t \in \mathbb{R}} \frac{f(a + te_j) - f(a)}{t}$$

For $f : A(\subseteq \mathbb{R}^n) \to \mathbb{R}^m$ differentiable at a in the interior of A, and d some unit vector in $\mathbb{R}^n$, the directional derivative of f at a in the direction of d, denoted $D_d f(a)$, is

$$\lim_{t \to 0, t \in \mathbb{R}} \frac{f(a + td) - f(a)}{t}$$

Jacobian Matrix For $f : A(\subseteq \mathbb{R}^n) \to \mathbb{R}^m$ differentiable at a in the interior of A, $\forall i = 1, \dots, m$ and $\forall j = 1, \dots, n$, $(D_j f_i)(a)$ is the component of $f'(a)$ in the i^{th} row and the j^{th} column.

$$D_j f_i(a) = \lim_{t \to 0, t \in \mathbb{R}} \frac{f_i(a + te_j) - f_i(a)}{t}$$

Extreme Values Always check:

- Points of discontinuity
- Points where the derivative doesn't exist
- Points where the derivative is 0
- Endpoints

For $f : \mathbb{R}^n \to \mathbb{R}^m$, the Hessian matrix is

$$H_f(a) = \begin{bmatrix} D_{11} f(a) & \cdots & D_{1n} f(a) \\ \vdots & & \vdots \\ D_{n1} f(a) & \cdots & D_{nn} f(a) \end{bmatrix}_{n \times n}$$

For $f : \mathbb{R}^2 \to \mathbb{R}$, if $H_f = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ is

- positive definite ($a > 0$, $\det H_f > 0$) $\Rightarrow$ local minimum
- negative definite ($a < 0$, $\det H_f > 0$) $\Rightarrow$ local maximum
- indefinite ($\det H_f < 0$) $\Rightarrow$ saddle point

If $\det H_f < 0$, there is not enough information available.

Quadratic Approximation The best quadratic approximation (really just a new version of Taylor) of f near a is

$$f(a + h) \approx f(a) + Df_a(h) + \frac{1}{2} h^T H_f(a) h$$

where

$$h^T H_f(a) h = \begin{bmatrix} h_1 & \cdots \end{bmatrix} \begin{bmatrix} H_f a \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \end{bmatrix} \in \mathbb{R}$$

Easy Inverse of a 2 × 2 Matrix For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Level Curves & Sets A **level set** of f is a set of the following form: $\{x \in A : f(x) = c\}$ for some constant c. If $n = 2$, then the level set is called a **level curve**.
If p is a critical point of f restricted to L, then for all tangent vectors d on L, the directional derivative of f in the direction of d has to be 0: $D_d f(p) = \langle \nabla f(p), d \rangle = 0$

Gradient The gradient of f at a is

$$\nabla f(a) = (D_1 f(a), \dots, D_n f(a))$$

It is the vector corresponding to the Jacobian matrix

$$f'(a) = \begin{bmatrix} D_1 f(a) & \cdots & D_n f(a) \end{bmatrix}$$

$\nabla f(a)$ tells the direction of greatest increase - the unit vector in that direction is

$$\frac{\nabla f(a)}{|\nabla f(a)|}$$

It is the direction of the **integral curve** and orthogonal to the **level set**.

Inverse Images Let $f : A \to B$, $C \subset B$. $f^{-1}(C)$ is this inverse image of C under f: $\{x \in A : f(x) \in C\}$. For this definition, f need not be invertible.
For A an open subset of $\mathbb{R}^n$ and $f : A \to \mathbb{R}^m$ continuous everywhere, for all subsets $W \subset \mathbb{R}^m$, $f^{-1}(W)$ is open.

Inverse Function Theorem Let A be an open subset of $\mathbb{R}^n$, $g : A \to \mathbb{R}^n$, $a \in A$, g continuously differentiable on some open ball containing a, and $g'(a)$ an invertible matrix.
Then there exist open subsets $V \subseteq A$, $W \subseteq \mathbb{R}^n$, a point $a \in V$, and $g(V) \subseteq W$ (the restriction of g to V) such that W is invertible with differentiable inverse, and $(Dg^{-1})_{g(x)} = (Dg_x)^{-1}$ for all $x \in V$.

Implicit Function Theorem Let $m \leq n$, A an open subset of $\mathbb{R}^n$, $f : a \to \mathbb{R}^m$ continuously differentiable everywhere, $L = \{x \in A : f(x) = 0\}$, $p \in L$, write $f'(p) = \begin{bmatrix} M & N \end{bmatrix}$, where M is an $m \times (n - m)$ submatrix, and N is an $m \times m$ submatrix, $p = (a, b)$ where $a \in \mathbb{R}^{n-m}$, $b \in \mathbb{R}^m$.
Then near p, L is the graph of a function. More specifically, there exist an open subset $W \subseteq A$ and an open set $V \in \mathbb{R}^{n-m}$ such that $p \in W$, $a \in V$, and there exists a function $\varphi : V \to W$ that is differentiable at a and $\varphi'(a) = -N^{-1}M$. Also, $\varphi(a) = p = (a, b)$.

Big Awkward Theorem Let $m \leq n$, $f : A \to \mathbb{R}^m$ continuously differentiable on an open set containing $p \in A$. Assume $f(p) = 0$, and that the largest invertible submatrix of $f'(p)$ has size $m \times m$.
Then there exists an open set $V \subseteq \mathbb{R}^n$ and a differentiable function $h : V \to \mathbb{R}^n$ with differentiable inverse such that for $(x_1, \dots, x_n) \in V$, $f \circ h(x_1, \dots, x_n) = (x_{n-m+1}, \dots, x_n)$, and the range of h contains an open set containing p.

Lagrange! Let A be a subset of $\mathbb{R}^n$, $f : A \to \mathbb{R}$ differentiable, $g : A \to \mathbb{R}^m$ continuously differentiable, $L = \{x \in A : g(x) = 0\}$. Suppose $p \in L$ is an extreme point for f restricted to L. Assume $g'(p)$ has m linearly independent rows/columns ("rank m").
Then $\nabla f(p) = \sum_{i=1}^m \lambda_i \nabla g_i(p) = \begin{bmatrix} \lambda_1 \cdots \lambda_m \end{bmatrix} g'(p)$ and $g(p) = 0$.

Other Stuff to Remember Determinants can give volume, but don't forget to take absolute value.
"If you can avoid doing hard and possibly meaningless work . . . "
When in doubt, add a clever 0, multiply by a clever 1, or compose by a clever identity.