Review Sheet for Final Exam

Standard disclaimer: The following represents a sincere effort to help you prepare for our exam. It is not guaranteed to be perfect. There might well be minor errors or (especially) omissions. These will not, however, absolve you of the responsibility to be fully prepared for the exam. If you suspect a problem with this review sheet, please bring it to my attention.

Time and place: the exam will take place Tuesday, December 17 from 4:15-6:15 in DBRT 117 (our MWF classroom). I'll hold a review session on Monday, December 16 from 7-9 PM in Hayes-Healy 125 (across from my office).

Format of the exam The format will be similar to that of the midterm, except that there will be no problem to complete ahead of time. Content-wise, the exam will be comprehensive with a strong tilt toward material covered since the midterm. New material includes the following.

- **Definitions and Statements** Here is a brief glossary of definitions and statements since the midterm that I will expect you to know.
  - **Definition** Let $F : X \to X$ be a mapping on some subset $X \subset \mathbf{R}^n$. Then F is a *contraction mapping* if there exists $C < 1$ such that
  $$\|f(x) - f(y)\| \leq C\,\|x - y\|$$
  for all $x, y \in X$.
  - **Contraction Mapping Theorem** Given a contraction mapping $F : X \to X$ on a closed subset $X \subset \mathbf{R}^n$, there exists a unique point $x_0 \in X$ such that $F(x_0) = x_0$. If, moreover, $x \in X$ is any other point, then $\lim_{n\to\infty} F^n(x) = x_0$.
  - **Existence and Uniqueness Theorem.** Let $U \subset \mathbf{R}^2$ be open and $F : U \to \mathbf{R}$ be a continuous function such that $\frac{\partial F}{\partial y}(t, y)$ exists and is continuous on U. Then for any $(t_0, y_0) \in U$, there exists an open interval $I \subset \mathbf{R}$ about t_0 and a unique differentiable function $y : I \to \mathbf{R}$ satisfying
    - $y(t_0) = y_0$, and
    - $y'(t) = F(t, y(t))$ for all $t \in I$.
  - **Inverse Function Theorem.** Let $U \subset \mathbf{R}^n$ be open and $F : U \to \mathbf{R}^n$ be a C^1 mapping. Suppose that $a \in U$ is a point such that the derivative map $DF(a)$ is invertible. Then there exists neighborhoods V of a and W of $f(a)$ such that $f : V \to W$ is invertible and the inverse function $f^{-1} : W \to V$ is C^1.
  - **Implicit Function Theorem.** Let $U \subset \mathbf{R}^n$ be open and $f : U \to \mathbf{R}$ be a C^1 function. Suppose that $a \in U$ is a point at which the partial derivative $D_j f(a)$ is non-zero. Then there exist neighborhoods $V \subset \mathbf{R}^n$ of a, $V' \subset \mathbf{R}^{n-1}$ of a_j' and a C^1 function $g : V' \to \mathbf{R}$ such that for any $x \in V$ we have $f(x) = 0$ if and only if $x_j = g(x_j')$.
- **Definition.** Given a set $X \subset \mathbf{R}^n$ and a point $a \in X$, we say that X is a *hypersurface near a* if there exists $j \in \{1, \ldots, n\}$, neighborhoods U of a and U' of a_j', and a C^1 function $g : U' \to \mathbf{R}^n$ such that $x \in X \cap U$ if and only if $x_j = g(x_j')$.

- **ODE skills** Finding solutions of 1st order linear and/or separable ODEs/initial value problems. Determining stable and unstable equilibrium solutions for 1st order autonomous ODEs. Determining asymptotic behavior of solutions to 1st order autonomous ODEs.
- **Other skills** Using Newton's method to approximate solutions of non-linear systems. Determining applicability of inverse/implicit function theorems. Determining when/where a given level set is a hypersurface. Determining whether a given vector is tangent/normal to a hypersurface at some point. Finding the intrinsic gradient of a smooth function restricted to a hypersurface. Finding the intrinsic critical points and determining maximum/minimum values of a smooth function along a hypersurface (method of Lagrange multipliers).

1. Hindsight: linear approximation

Broadly speaking, this course has been about differentiation and it's applications in the context of multivariable functions and mappings. The main idea behind differentiation is that a mathematical object can often be understood better if, on a small scale, it resembles something linear. We pursued this idea analytically (i.e. for functions/mappings), algebraically (for systems of equations), and finally geometrically (for subsets of $\mathbf{R}^n$).

1.1. **The analytic point of view.** A mapping $f : U \subset \mathbf{R}^n \to \mathbf{R}^m$ is *differentiable* at a point $a \in U$ if for x near a, the difference $f(x) - f(a)$ is well-approximated by a linear function of $x - a$. More precisely, $f(x) = f(a) + Jf(a)(x - a) + E(x - a)$ where $Jf(a)$ is the Jacobian matrix of f at a and $E(h)$ is 'small enough' in the sense that $\lim_{\mathbf{h}\to 0} \|E(\mathbf{h})\| / \|\mathbf{h}\| = 0$.

1.2. **The algebraic point of view.** One cannot usually solve a random system of m equations in n real variables. However, one can often *approximate* solutions to the system if, on a small scale, the given system closely resembles some solveable *linear* system of equations. More precisely, one writes the given system in the form $F(x) = b$, for some mapping $F : U \subset \mathbf{R}^n \to \mathbf{R}^m$. Then if we have a solution $f(x_0) = b_0$ for some b_0 close to b, and if F is differentiable at b_0, the given system is well-approximated by the linear system $JF(x_0)(x - x_0) = b - b_0$. If all goes well, then the solution $x = x_1$ of this linear system will be quite close to a solution of the given non-linear system. And the really great thing is that one can then repeat this process, substituting x_1 for x_0 and $F(x_1)$ for $b_0 = F(x_0)$ to further improve the approximate solution. This algorithm is known as Newton's method.

There are (at least) two important theoretical consequences of these ideas. The *inverse function theorem* concerns systems with the same number of equations as unknowns and deals with the prospect of solving the same system $f(x) = b$ for various values of b. The *implicit function theorem* concerns systems $f(x) = b$ with more equations than unknowns and deals, but rather than trying to solve the system for different values of b, one fixes b and tries to 'parametrize' the set of solutions x by solving for some coordinates of x in terms of others. In class, we only stated and discussed the implicit function for *one* equation of many variables, and in this case one seeks to solve $f(x) = b$ for *one* variable in terms of the others.

1.3. **The geometric point of view.** A subset $X \subset \mathbf{R}^n$ is a manifold of dimension k if, when we look at X under a microscope near any point $a \in X$, then $X - a$ (i.e. X shifted by $-a$ so that the point of interest becomes the origin in $\mathbf{R}^n$) is nearly indistinguishable from a k-dimensional linear subspace (which we call T_aX) of R^n. We only pursued this idea in two special cases: when $k = \dim X = 1$, X is a curve; and more recently and at greater length, when $\dim X = n - 1$, X is a smooth hypersurface. As it turns out there are three different, but ultimately equivalent, ways of describing smooth hypersurfaces (all three generalize to other dimensions k.

- X **is locally a smooth graph.** We have a C^1 function $g : U' \subset \mathbf{R}^{n-1} \to \mathbf{R}$ and a coordinate index $j \in \{1, \dots, n\}$ such that points x near a belong to X if and only if $x_j = g(x'_j)$. In this case, the tangent space T_aX is just the graph of the linearization of g; i.e. $\mathbf{v} := x - a \in T_aX$ if and only if $x_j = a_j + Jg(a'_j)(x_j - a_j)$.

  Note that in order to make visual comparison easier, one usually translates T_aX in pictures so that it passes through a rather than 0. This translated version '$a + T_aX$' of T_aX is known as an *affine subspace* of $\mathbf{R}^n$.
- X **is locally a smooth level set.** We have a C^1 function $\rho : V \subset \mathbf{R}^n \to \mathbf{R}$ such that $x \in X \cap V$ if and only if $\rho(x) = 0$. It is important to further impose the 'non-degeneracy condition' $\nabla\rho(x) \neq 0$ for $x \in X \cap V$. The significance of this condition becomes apparent when we use the defining function ρ to describe the tangent space T_aX of X at a: geometrically, a vector $\mathbf{v} := x - a$ belongs to T_pX if and only if $0 = \nabla\rho(a) \cdot \mathbf{v}$. In function-speak, $\mathbf{v}$ belongs to the kernel of the linear transformation $Df(a)$. In calculus-speak, $\mathbf{v} \in T_aX$ means that x is is in the level set $0 = a + \nabla\rho(a) \cdot (x - a)$ of the linear approximation of ρ.
- X **is locally smoothly parametrizable.** We have an open set $V \subset X$ and an injective C^1 function $\Psi : V' \subset \mathbf{R}^{n-1} \to U$ whose image $\Psi(V')$ is $X \cap V$. If $\Psi(x') = x \in X$, then it is common to refer to x' as the *coordinates* of x (relative to the parametrization Ψ). Here again, there is an important and additional non-degeneracy condition: the Jacobian matrix $J\Psi(x')$ must have full rank at each point $x' \in V'$. This is equivalent to saying that the $n - 1$ columns of $J\Psi(x')$ are independent, or that the linear transformation $D\Psi(x') : \mathbf{R}^{n-1} \to \mathbf{R}^n$ is injective.

  If $a = \Psi(a')$, then T_aX is just the column space of $J\Psi(a)$ (i.e. the range of $D\Psi(a)$). Alternatively, the translate $a + T_aX$ is parametrized by the linear approximation $\Psi(a') + D\Psi(a')(x' - a)$ of Ψ.

Formally, one verifies that the analytic, algebraic and geometric points of view are consistent with one another by judiciously applying the inverse/implicit function theorems and the chain rule.

It takes some time to get used to all these equivalent formulations concerning hypersurfaces and their tangent spaces and to move back and forth gracefully from one point of view to the other. But all three points of view are valuable, and all become very intuitive with time, patience, and effort.