Math 5610—Fall 2017—Review and Summary

Technicalities:

- **Final Exam.** Thursday, December 14, 2017, 10:30am-12:30pm, in LCB 215.
- The exam is closed books and notes, no electronic devices. You can use your own paper.
- You’ll be asked to answer 8 out of 13 questions. Clearly indicate which questions you want counted. You should have plenty of time.
- Accuracy is much more important than speed. You must write your answers and describe your figuring and reasoning in such a way that I can understand it. When you get the exam read all the questions and then decide which ones to answer. Answer the easiest first. On your way out pick up an answer sheet. I will send you an email with your score and your final grade as soon as I have it. (Clearly print your email address on the exam.)
- **To avoid distraction and confusion I am unable to answer questions during the exam.** If there is an exam question that you don’t understand answer another one. If you think there is something wrong or inconsistent with a question write a note and you will get credit. Answer another question as well, just to be sure.

- **Rules to live by**
  - Asking the right questions is more important than knowing some right answers.
  - If you don’t understand a question your won’t be able to answer it. Don’t even try. Figure out first what the question is asking.
  - Always have expectations. (If they are met, good. Otherwise find the error or figure out what it is that you did not understand.) Think about your expectations, and make them explicit, before starting to work on a problem.
  - Always check your answers.
  - Everybody makes mistakes. That’s OK. It’s not OK not to catch those mistakes.
  - To solve a difficult problem build a hierarchy of simpler problems leading up to it. Never hesitate to make simplifying assumptions. Worry about those assumptions only after you answer the simpler question.
  - First, assume all horses are spherical.
  - Even though a problem may be difficult and take a lot of time for its solution, once you solve it the next similar problem will be easier.
  - Match the number of parameters and the number of conditions.
  - Choose the right level of abstraction and generality. (Example: the interpolant of symmetric data is symmetric.)
  - Go back to the definition.
  - Focus on understanding rather than on recipes.
  - Understanding something means being able to explain it in simpler terms and relating it in various ways to other things.
  - Make sure there is a great deal of redundancy in your understanding.
  - To succeed in mathematics you must master its language.
  - If something has a lot of names it must be important.

---

1. \TeX\ processed by Peter Alfeld November 30, 2017. JWB 127, 801-581-6842, pa@math.utah.edu.
2. This phrase was invented by John Lund, a student in an earlier incarnation of this class. He volunteered to draw up a list (which overlapped with this list, but was different) and present it to that class. His main occupation, however, was serving as an officer on a US nuclear submarine.
Quotes

• The first thing I ask when I find myself in a room full of manure is “where is the pony”. (Ronald Reagan)
• The purpose of computing is insight, not numbers (Hamming.)
• It ain’t so much the things we don’t know that get us into trouble. It’s the things we do know that just ain’t so. (Artemus Ward.)
• In theory, theory and praxis are the same. In praxis, they aren’t. (Richard Nixon.)
• Don’t worry about your difficulties with mathematics. I can assure you that mine are still greater. (Albert Einstein, to his neighbor’s young daughter.)
• A Vulgar Mechanick can practice what he has been taught or seen done, but if he is in an error he knows not how to find it out and correct it, and if you put him out of his road, he is at a stand; Whereas he that is able to reason nimibly and judiciously about figure, force and motion, is never at rest till he gets over every rub. (Isaac Newton)

The Subject

• Linear Algebra. The algebra of linear functions between finite dimensional vector spaces.
• “Matrix” is a synonym for “linear function”
• Matrix multiplication = function composition.
• Orthogonal Matrices. \((A^{-1} = A^T)\) Columns form an orthonormal basis. They don’t amplify errors.
• vector and matrix norms, induced matrix norm (or 'operator norm')
• Linear System. \(Ax = b\)
  • \(A = LU\) + forward and backward substitution
• Never invert a matrix! It’s unnecessary, it’s expensive, it’s likely to introduce additional round-off errors, and it destroys sparsity.
• Backward Error Analysis, condition number
  \[
  A(x - e) = b - r
  \]
  \[
  \frac{1}{\|A\|\|A^{-1}\|} \|r\| \leq \|e\| \leq \|A\|\|A^{-1}\| \|r\|
  \]
  loose approximately \(\log_{10} \|A\|\|A^{-1}\|\) significant digits.
• Special Case: A symmetric matrix \(A\) is positive definite if \(x^T Ax > 0\) for all \(x \neq 0\). Positive definite systems arise in minimization problems. Use the Cholesky decomposition: \(A = LL^T\). This does not require pivoting.
  • Least Squares: \(\|Ax - b\|_2 = \text{min} \).
    — Normal equations \(A^T Ax = A^T b\). Conceptually simple but numerically inferior.
    — QR factorization approach is better.
• Ill conditioning arises naturally and is incurable once it’s there. Use a clever basis to avoid it.
• QR factorization. Compute via Householder Reflections
  \[
  H = I - 2uu^T, \quad \|u\|_2 = 1.
  \]
• Singular Value Decomposition, SVD: The last resort in difficult circumstances, and useful for theoretical analysis. It can be used to analyze linear systems and least squares problems, determine the rank, determinant, norm and condition number of a matrix, and it can also be used for image compression.
  \[
  A = U\Sigma V^T
  \]
  \(U\) and \(V\) are orthogonal, \(\Sigma\) is diagonal. Singular values. Left and right singular vectors.
• **Eigenvalue Problems** $Ax = \lambda x$. This is a nonlinear system. eigenvalues, eigenvectors, eigenvectors only determined up to a constant.

• **Eigenvalues** are roots of the characteristic polynomial

\[ p(\lambda) = \det(A - \lambda I) \]  

(4)

• For any polynomial $p$ there is a matrix that has $p$ as its characteristic polynomial, e.g., the **companion matrix**.

• **Gershgorin Theorem:** Suppose $Ax = \lambda x$. Then, for some $i$,

\[ |a_{ii} - \lambda| \leq \sum_{j \neq i} |a_{ij}|. \]  

(5)

• To find **roots of a polynomial** find the eigenvalues of its companion matrix.

• **Power Method:** find one eigenvalue and corresponding eigenvector. Shift of Origin and inverse iteration.

• **QR algorithm:** find all eigenvalues (and corresponding eigenvectors). Key ingredients include: simultaneous orthogonal iteration, use of orthogonal matrices, Hessenberg matrices, maintaining $O(n^2)$ effort at all times, decoupling, shifts of origin, executing complex double shifts in real arithmetic. The most complex algorithm we encountered this semester.

• **Linear Programming:** Minimize or maximize a linear functions subject to linear constraints. Various forms of an LP problem. Discrete $L_1$ and $L_\infty$ approximation, transportation problem.

• **Simplex Method.** Go from vertex to vertex.

• **Polynomials.** Evaluate by Horner’s Scheme (nested multiplication, synthetic division) and its generalizations. The space of polynomials is closed under multiplication, addition, composition, integration, and differentiation. Degree, exact degree, degree no larger than, constant, linear, quadratic, cubic, quartic, quintic.

• **Fixed Point Iteration.** Find a function $g$ such that $f(x) = 0$ is equivalent to $x = g(x)$. Then start with some $x_0$ and iterate:

\[ x_{k+1} = g(x_k) \quad k = 0, 1, 2, \ldots \]  

(6)

• **Newtons’s Method.** A special case of a fixed point iteration. Basic idea: Linearize and Iterate. That idea has a wide range of applications. For $f(x) = 0$ it becomes

\[ x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \]  

(7)

• The choice of **starting value** depends on the problem.

• The fixed point iteration (6) converges of order $p$ to $\alpha$ if

\[ g(\alpha) = \alpha, \quad g'(\alpha) = g''(\alpha) = \ldots = g^{(p-1)}(\alpha) = 0, \quad g^{(p)}(\alpha) \neq 0. \]  

(8)

• **Aitken Acceleration.** Make a linearly convergent sequence $p_0, p_1, p_2, \ldots$ into a quadratically convergent sequence $\hat{p}_0, \hat{p}_1, \hat{p}_2, \ldots$

\[ \hat{p}_n = p_n - \frac{(p_{n+1} - p_n)^2}{p_{n+2} - 2p_{n+1} + p_n}, \quad n = 0, 1, 2, \ldots \]  

(9)

• **Interpolation** means exact reproduction of data.

• **Lagrange interpolation** = interpolation to function values by a polynomial.
  – **power form**, Vandermonde Matrix, Existence, Uniqueness
  – **Newton Form.** Add one point at a time.
  – **Lagrange Form.** A special case of cardinal form.
• **Interpolation Error:** Let $p$ be the polynomial interpolating to $f$ at $n + 1$ distinct points $x_0, \ldots, x_n$. Then
\[
f(x) - p(x) = \frac{\prod_{i=0}^{n} (x - x_i)}{(n + 1)!} f^{(n+1)}(\xi).
\] (10)

• **The Runge Phenomenon.** Interpolation by a high degree polynomial can lead to large oscillations (and errors).

• **Hermite interpolation** = interpolation to function and derivative values by a polynomial.

• **Cardinal Form of an Interpolant** Use the function (and derivative) values as coefficients.

• **Numerical Integration** Differentiate the interpolant.

• **Newton Cotes Formulas** Open, closed, simple, composite. Interpolate at equally spaced points.

• **Gaussian Quadrature.** Use weights and knots to obtain exactness for polynomials of degree as high as possible. Gaussian genius allowed solution of the nonlinear system.

• **Numerical Differentiation.** It’s difficult or impossible. However, if it can’t be helped differentiate the interpolant. Choose the knots judiciously.

• **Symmetry.** Abandon it only if you have to.

• **Method of Undetermined Coefficients.** Write down the general form of a method. Find the coefficients by requiring exactness for specific functions.

• **Piecewise Polynomial Functions.**
  – **Piecewise linear:** Simple, but effective, particularly in Computer Graphics which is ultimately pixel based.
  – **Piecewise Cubic Hermite:** Once differentiable, interpolate to function values and derivatives.
  – **Cubic Splines:** twice differentiable, interpolate to function values only. Two end conditions must be imposed: forced end condition, natural end condition, not-a-knot condition.

• **The Weierstrass Theorem.** Polynomials can be used to approximate continuous functions on a closed and bounded interval arbitrarily well. Constructive proof by Bernstein.

• **The Bernstein Bézier Form of a polynomial.** Address algebraic questions, such as shape or smoothness, geometrically.

• **Bézier ordinates, domain points, control points, control polygon, subdivision, de Casteljau Algorithm.**

• **Spline spaces:** smooth piecewise polynomial functions. In more than one variable, the dimension (and other attributes) depend not just on the topology, but also on the geometry, of the underlying partition.

• **Approximation of Functions.** Minimize the norm of the difference of a function and a linear combination of basis functions. The norm can be the 2 norm, or some other norm.

• **Basis Functions.** Pick them according to your problem. Use them to incorporate singularities (explosions), exponential growth or decay, or periodicity, for example.

• **Orthonormal Basis.** Many kinds of orthogonal polynomials.

• **Fourier Series.** Approximate periodic functions a series of sin and cos terms.

• **Fast Fourier Transform.** FFT, compute a discrete Fourier transform in $O(N \log N)$ instead of $O(N^2)$ operations.

• **The Gibbs Phenomenon.** Discontinuities in the approximated function cause oscillations in the error whose amplitudes do not diminish as the number of terms increases.

• **Wavelets.** Multiresolution analysis. Resolution should be appropriate for one’s purposes, and local phenomena should have local effects.

• **Padé Approximants** of a function $f$ are rational functions whose Taylor Series match the first few terms of the Taylor Series of $f$.

**Sample Exam Questions**
-1- (The Gershgorin Theorem.) The matrix
\[
A = \begin{bmatrix}
-5 & 1 \\
4 & 7 \\
\end{bmatrix}
\] (11)
has two real eigenvalues. Use the Gershgorin Theorem to find two intervals that contain them. Can you improve this result by observing that the eigenvalues of \(A\) are also those of \(A^T\)?

-2- (\(L_1\)-approximation.) What is the linear function
\[
l(x) = \alpha x + \beta
\] (12)
that minimizes
\[
\int_0^1 |\sin x - l(x)| \, dx?
\] (13)

-3- (Method of Undetermined Coefficients.) Consider the (somewhat peculiar) quadrature rule
\[
\int_0^{1} f(x) \, dx = w_1 f(0) + w_2 f\left(\frac{1}{4}\right) + w_3 f(1) + E.
\] (14)
Find the values of \(w_1, w_2,\) and \(w_3\) for which the rule is exact (i.e., \(E = 0\)) for polynomials of degree as high as possible. What is that degree?

-4- (True or False.) A linear system of \(m\) equations in \(n\) unknowns is square if \(m = n\), overdetermined if \(m > n\), and underdetermined if \(m < n\). Indicate whether the following statements are true or false:

a. For a square linear system, all of the following cases are possible: There is no solution, 1 solution, or infinitely many solutions. These are all possibilities.

b. A square linear system may have precisely 15 different solutions.

c. An underdetermined linear system may have no solution.

d. An underdetermined linear system may have a unique solution.

e. An underdetermined linear system may have infinitely many solutions.

f. An overdetermined linear system cannot have a solution.

g. An overdetermined linear system cannot have more than one solution.

-5- (Condition Number.) Define the term “condition number” of a square matrix with respect to
\[
\|A\| = \max_{\|x\|=1} \|Ax\|_2.
\] (15)
What is the condition number of
a) a 1 \times 1 matrix?

b) a diagonal matrix?

c) an orthogonal matrix?

In each case give reasons for your answer.

-6- (Gaussian Elimination.) A square matrix \(A\) with entries \(a_{ij}\) is called a lower Hessenberg matrix if \(a = 0\) whenever \(j > i + 1\).

a) Draw a schematic picture of a lower Hessenberg matrix indicating the zero and the non-zero parts.

b) Describe an efficient variant of Gaussian Elimination to solve \(Ax = b\) where \(A\) is a lower Hessenberg matrix (Don’t go into detail, just point out any key differences to ordinary Gaussian Elimination). Why would you not use the general purpose version of Gaussian elimination that we discussed in class? (For simplicity, ignore pivoting issues.)
-7- (Fixed Point Iteration.) Consider the fixed point iteration

\[ x_0 = 1, \quad x_{k+1} = \sin x_k, \quad k = 0, 1, 2, \ldots \]  

(Naturally, the angle is measured in radians.) Show that this iteration converges. What does it converge to? How fast? What do you expect to happen if the iteration is actually run on a computer? (You are probably on the wrong track if you carry out more than 20 or so steps on your calculator.)

-8- (Newton's Method, once again.) Describe a class of functions \( f \) for which Newton's method

\[ x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \]  

converges of order 3 (but not of order 4) to a root \( \alpha \) of

\[ f(x) = 0 \]  

(presuming of course, that it converges at all).

-9- (The Singular Value Decomposition.) Suppose \( A \) is an \( m \times 1 \) matrix:

\[ A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}. \]  

Describe the Singular Value Decomposition of \( A \).

-10- (Householder Reflections.) Define what a Householder Reflection is, and show that it is symmetric and orthogonal. Give an example of an orthogonal matrix which is not symmetric (and hence not a Householder Reflection).

-11- (Continuous Linear Least Squares.) Show that the Hilbert Matrix is positive definite.

-12- (Backward Error Analysis.) Suppose you are solving the square linear system \( Ax = b \) and you are concerned about the effects of perturbations of the coefficient matrix \( A \). (In class we discussed the effects of perturbations of \( b \).) Suppose you solve the perturbed system

\[ (A + E)(x + e) = b \]  

where \( E \) is the perturbation of \( A \) and \( e \) is the resulting error in \( x \). Show that (for any vector norm and the induced matrix norm)

\[ \frac{\|e\|}{\|x + e\|} \leq \|A\|\|A^{-1}\|\|E\|/\|A\|. \]  

(In other words ill conditioning affects errors in the matrix similarly as errors in the right hand side.)
\[ f(x) = 0 \implies f(x) = 0 \]

\( x_0 \) given

\[ x_{n+1} = g(x_n) \]

\[ \alpha - x_{n+1} = O\left((\alpha - x_n)^{\rho}\right) \]

\[ g(x) = \alpha \; g'(x) = \ldots = g^{(k)}(x) = 0 \]

\[ g^{(k+1)}(\alpha) \neq 0 \]

**Inverse Interpolation**

\[ f(x) = 0 \quad x = f^{-1}(0) \]

\[ x_i \; f(x_i) \]

\[ \cdot \]

\[ \cdot \]