Wed. Dec 5

- slides from prof. Alberts' lecture are posted on lecture page.
- pretend quiz today (if you want)
- appendix with proof of Spectral Theorem in today's notes

Math 2270-002
just Final Exam Review Information

Our final exam is next Wednesday afternoon, December 12, 1:00-3:00 p.m., in our usual MWF classroom LCB 215. I will let you work until 3:30 p.m. if you wish. There will be a review session Monday December 10, from 1:00-3:00 p.m., with room TBA. Most of that session will be devoted to going over a practice exam (which I will post by later this week), but please bring any other questions you may have.

The exam will be comprehensive. Precisely, you can expect anything we covered from sections 1.1-1.9, 2.1-2.3, 3.1-3.3, 4.1-4.7, 4.9, 5.1-5.5, 6.1-6.8, 7.1-7.2 as well as some supplementary material - we'll discuss in more detail below. In addition to being able to do computations, you should know key definitions, the statements of the main theorems, and why they are true. The exam will be a mixture of computational and conceptual questions. As on the midterms, I will test conceptual understanding with true-false and "example" questions, and by asking you to explain your reasoning at various places. The exam is closed book and closed note, except for a single index card of size up to 5 by 7 inches. You may use a scientific calculator to check arithmetic, if you wish. or 6 by 8

Exam material will be weighted towards topics which have not yet been tested, i.e. Chapter 6-7 material (which relies on the earlier chapters).

Copies of my final exams from previous years can be found on my "old classes" web page, although we used a different textbook and had somewhat different emphases back then. This semester's midterms, quizzes, homework, class notes, and the text are good references. It always worked well for me as a student to make my own course outline with the key ideas (which I would then make sure I could explain and work with).

revisit day 1 – syllabus

Learning Objectives for 2270

Computation vs. Theory: This course is a combination of computational mathematics and theoretical mathematics. By theoretical mathematics, I mean abstract definitions and theorems, instead of calculations. The computational aspects of the course may feel more familiar and easier to grasp, but I urge you to focus on the theoretical aspects of the subject. Linear algebra is a tool that is heavily used in mathematics, engineering, science and computer science, so it will likely be relevant to you later in your career. When this time comes, you will find that the computations of linear algebra can easily be done by computing systems such as Matlab, Maple, Mathematica or Wolfram alpha. But to understand the significance of these computations, a person must understand the theory of linear algebra. Understanding abstract mathematics is something that comes with practice, and takes more time than repeating a calculation. When you encounter an abstract concept in lecture and the text, I encourage you to pause and give yourself some time to think about it. Try to give examples of the concept, and think about what the concept is good for.

The essential topics

Be able to find the solution set to linear systems of equations systematically, using row reduction techniques and reduced row echelon form - by hand for smaller systems and using technology for larger ones. Be able to solve (linear combination) vector equations using the same methods, as both concepts are united by the common matrix equation $A\mathbf{x} = \mathbf{b}$.

@ wikipedia "linear algebra"

Be able to use the correspondence between matrices and linear transformations - first for transformations between $\mathbb{R}^n$ and $\mathbb{R}^m$, and later for transformations between arbitrary vector spaces.

Become fluent in matrix algebra techniques built out of matrix addition and multiplication, in order to solve matrix equations.

Understand the algebra and geometry of determinants so that you can compute determinants, with applications to matrix inverses and to oriented volume expansion factors for linear transformations.

wiki starts here

Become fluent in the language and concepts related to general vector spaces: linear independence, span, basis, dimension, and rank for linear transformations. Understand how change of basis in the domain and range effect the matrix of a linear transformation.

Be able to find eigenvalues and eigenvectors for square matrices. Apply these matrix algebra concepts and matrix diagonalization to understand the geometry of linear transformations and certain discrete dynamical systems. — we did Markov chains with stochastic matrices (& google)

Understand how orthogonality and angles in $\mathbf{R}^2, \mathbf{R}^3$ generalize via the dot product to $\mathbf{R}^n$, and via general inner products to other vector spaces. Be able to use orthogonal projections and the Gram-Schmidt process, with applications to least squares problems and to function vector spaces.

Know the spectral theorem for symmetric matrices and be able to find their diagonalizations. Relate this to quadratic forms, ~~constrained optimization problems~~, and to the ~~singular value decomposition for matrices~~. Learn some applications to image processing and/or statistics.

Week-by-Week Topics Plan

Topic schedule is subject to slight modifications as the course progresses, but exam dates are fixed.

Week 1: 1.1-1.3; systems of linear equations, row reduction and echelon forms, vector equations.

Week 2: 1.3-1.5; matrix equations, solution sets of linear systems, applications.

Week 3: 1.6-1.8; applications, linear dependence and independence, linear transformations and matrices.

Week 4: 1.8-1.9, 2.1-2.2; linear transformations of the plane, introduction to matrix algebra.

Week 5: 2.3-2.5; matrix inverses, partitioned matrices and matrix factorizations. 2.4 skipped 2.4-2.5 except col-row "outer product" mult. did do that (2.4)

Week 6: 3.1-3.3; determinants, algebraic and geometric properties and interpretations. **Midterm exam 1 on Friday September 28** covering material from weeks 1-6.

Week 7: 4.1-4.3; vector spaces and subspaces, nullspaces and column spaces, general linear transformations.

Week 8: 4.3-4.6; linearly independent sets, bases, coordinate systems, dimension and rank

Week 9: 4.6-4.7, 5.1-5.2; change of basis, eigenvectors and eigenvalues, and how to find them. § 4.9 Markov.

Week 10: 5.3-5.5; diagonalization, eigenvectors and linear transformations, complex eigendata.

Week 11: ~~5.6~~, 6.1-6.2; ~~discrete dynamical systems~~, introduction to orthogonality. **Midterm exam 2 on Friday November 9** covering material from weeks 7-11

Week 12: 6.3-6.5; orthogonal projections, Gram-Schmidt process, least squares solutions

Week 13: 6.6-6.8; applications to linear models; inner product spaces and applications with Fourier series.

Week 14: 7.1-~~7.3~~ 7.2; diagonalization of symmetric matrices, quadratic forms, ~~constrained and unconstrained optimization~~.

Week 15: ~~7.4; singular value decomposition~~, applications, course review.

Week16: **Final exam Wednesday December 12, 1:00 - 3:00 p.m. in classroom ~~WEB ST 208~~ LCB 215 (here!). This is the University scheduled time.**

for sure likely

Topics/concepts list for final exam

Sections 1.1-1.9, 2.1-2.3, 3.1-3.3 the material on the first midterm:

The matrix equation $A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$ arises in a number of different contexts: it can represent a system of linear equations in the unknown $\underline{\boldsymbol{x}}$; a vector linear combination equation of the columns of A, with weights given by the entries of $\underline{\boldsymbol{x}}$; in the study of the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ given by $T(\underline{\boldsymbol{x}}) = A\,\underline{\boldsymbol{x}}$.

An essential tool is the reduced row echelon form of a matrix (augmented or unaugmented), and what it tells you about (1) solutions to matrix equations; (2) the structure of solution sets to matrix equations; (3) column dependencies of a matrix. (We turned the column dependency idea around to understand why each matrix can have only one reduced row echelon form.) We also discussed matrix transformations $T(x) = A\,x$ from $\mathbb{R}^n$ to $\mathbb{R}^m$, as a prelude to general linear transformations $T : V \to W$ that we discussed later in the course and that appear in subsequent courses. We focused on the geometry of linear transformations from $\mathbb{R}^2 \to \mathbb{R}^2$, $\mathbb{R}^2 \to \mathbb{R}^3$, $\mathbb{R}^3 \to \mathbb{R}^2$, etc., which are important in their own right, and where visualization helps develop intuition.

Square matrices and linear transformations from $\mathbb{R}^n \to \mathbb{R}^n$ are an important special case. Invertible matrices are useful in matrix algebra computations. One should know how to find matrix inverses and use them.

3.1-3.3 determinants.
how to compute via cofactor expansions or elementary row (or column) operations
$|A| \neq 0$ as a test for matrix invertibility, (or equivalently whether the columns of the matrix are a basis for $\mathbb{R}^n$, whether $rref(A) = I$, or anything else on the long list of equivalent characterizations).
$|det(A)|$ as area/volume expansion factor for $T(\boldsymbol{x}) = A\,\underline{x}$.
Adjoint formula for A^{-1}

....
material after the first midterm

4.1 vector spaces and sub vector spaces (subspaces) - abstract definitions.
realization of subspaces as null spaces or as spans of collections of vectors
how to check if a subset is a subspace.
examples such as polynomial vector spaces, matrix vector spaces, $\mathbb{R}^n$, and subspaces of all of these.

4.2 *Nul A* and *Col A* for $T(\underline{\boldsymbol{x}}) = A\,\underline{\boldsymbol{x}}$; *Kernel T* and *Range T* for general linear transformations $T : V \to W$
definition of linear transformation, examples.
how to find *Nul A* and *Col A*, and bases for each.

4.3 linearly independent/dependent sets; bases for vector spaces (including subspaces).
how to check whether the vectors in a set span a vector space.
how to check whether a set of vectors is linearly independent.
how to build up bases as growing sets of independent vectors, one vector at a time, until the set spans.
how to cull dependent vectors from a spanning set, until it is an independent set.

4.4 every basis of n vectors for a vector space V yields a coordinate system, via the coordinate isomorphism with $\mathbb{R}^n$.

answering questions about span and linear independence for sets of vectors in V by using coordinates with respect to a basis.

favorite examples include P_n, $M_{m \times n}$, the polynomial and matrix spaces.

4.5 dimension of a vector space. basic facts about dimension, number of vectors required to span, maximum number of independent vectors, dimensions of subspaces.

4.6 rank of a matrix. rank + nullity theorem.

connection to reduced row echelon form of the matrix.

how to find $Row\ A$, $Nul\ A^T$.

what $Nul\ A$, $Row\ A$, $Col\ A$, $Nul\ A^T$ have to do with the geometry of the transformation $T(\underline{\boldsymbol{x}}) = A\,\underline{x}$.

4.9 applications to Markov Chains with stochastic matrices; Google page rank vector.

5.1-5.3 eigenvectors, eigenvalues, diagonalization

characteristic polynomial to find eigenvalues of a matrix

$E_{\lambda = \lambda_i} = Nul\left(A - \lambda_i I\right)$. Finding eigenvector as weights for column dependencies, or the "old" way via backsolving.

diagonalizable and non-diagonalizable matrices.

Using $A = PDP^{-1}$ to compute large matrix powers.

improved understanding of the transformation $T(\underline{\boldsymbol{x}}) = A\,\underline{\boldsymbol{x}}$ in terms of $\mathbb{R}^n$ basis made out of eigenvectors, as compared to the standard basis.

5.4 matrix of a linear transformation $T : V \to W$.

Finding the matrix of a linear transformation, given bases for V, W

What the columns of such a matrix must be

special cases:

matrix of $T : \mathbb{R}^n \to \mathbb{R}^n$, $T(x) = A\,x$ with respect to an eigenbasis of A or a "better" basis than the standard one, and its relationship to the identity $B = P^{-1}AP$.

5.5 complex eigendata $\longrightarrow$ people who take 2280 (or higher-level physics) will see a lot of this

Appendix B the complex plane

material after the second midterm $\geq$ 50% of final exam

6.1 dot product (inner product), length, orthogonality.
algebra of dot product
Pythagorean Theorem
$proj_L\,\underline{\boldsymbol{w}}$ where $L = span\{\underline{\boldsymbol{v}}\}$ (or onto higher-dim'l subspace)
$\underline{\boldsymbol{z}} = \underline{\boldsymbol{w}} - proj_L\,\underline{\boldsymbol{w}}$ is $\perp$ to L.
angles in $\mathbb{R}^n$
orthogonal complements and how to find them
four fundamental subspaces of a matrix, revisited

6.2-6.3 orthogonal and ortho-normal sets in $\mathbb{R}^n$.
coordinates in a subspace with respect to an orthonormal (or orthogonal basis).
projection onto a subspace having an orthonormal (or orthogonal) basis.

6.4 Gram-Schmidt orthogonalization
the algorithm, algebraically and geometrically
A= QR factorization

6.5 Least-squares solutions to inconsistent systems
geometric meaning, and computation via orthonormal basis for *Col A*
alternate solution using normal equations, and why this method works
computing projections without having an orthonormal basis for the subspace.

6.6 least squares solutions for "linear models" via the normal equations
esp. best-line fit to a collection of data points

6.7-6.8 inner product spaces
magnitude, orthogonal, orthonormal, projection with orthogonal or orthonormal basis.
Fourier series (I will provide formulas for Fourier coefficients if necessary I ask such a question.)
lots of these in later math & physics courses

7.1 Spectral theorem for symmetric matrices
constructing orthonormal eigenbases for symmetric matrices
outer product method of multiplying matrices
spectral decomposition theorem

7.2 quadratic forms
expressed via a symmetric matrix
diagonalizing symmetric forms with orthogonal change of variables
applications to conics and quartic surfaces
positive definite and semi-definite quadratic forms.

fluency in the defintions and concepts
ability to create examples illustrating definitions and concepts
ability to discern whether statements are true or false, based on the material we've covered.

Appendix: Proof of the Spectral Theorem (matrix version). There are more general versions (see Wikipedia)

<u>Spectral Theorem</u> Let A be an $n \times n$ symmetric matrix with real number entries. Then all of the eigenvalues of A are real, and there exists an orthonormal eigenbasis $B = \{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \ldots \underline{\boldsymbol{u}}_n\}$ consisting of eigenvectors for A.

<u>part 1</u>: Let A be an $n \times n$ symmetric matrix with real number entries. Then all of its eigenvalues are real.

<u>proof</u>: Recall that for a matrix with real entries, if $\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}}$ is a complex eigenvector (with $\underline{\boldsymbol{u}}$, $\underline{\boldsymbol{v}}$ real vectors), with eigenvalue $a + b\,i$, then the complex conjugate $\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}}$ is an eigenvector with conjugate eigenvalue $a - b\,i$: (class notes November 6; text section 5.5). Consider the complex scalar

$$(\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}})^T A\,(\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}}).$$

Computing the product by computing the second product first yields

$$(\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}})^T [A\,(\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}})] = (\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}})^T (a + b\,i)(\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}}) = (a + b\,i)(\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}})^T(\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}})$$
$$= (a + b\,i)\left(\|\underline{\boldsymbol{u}}\|^2 + \|\underline{\boldsymbol{v}}\|^2\right).$$

Computing the triple product by computing the first product first yields

$$\left[(\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}})^T A\right](\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}}) = \left[(\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}})^T A^T\right](\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}}) \quad \text{(since } A \text{ is symmetric)}$$
$$= [A\,(\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}})]^T(\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}}) = (a - b\,i)\,(\underline{\boldsymbol{u}} - i\,\underline{\boldsymbol{v}})^T(\underline{\boldsymbol{u}} + i\,\underline{\boldsymbol{v}})$$
$$= (a - b\,i)\left(\|\underline{\boldsymbol{u}}\|^2 + \|\underline{\boldsymbol{v}}\|^2\right).$$

We deduce that $(a + b\,i) = (a - b\,i)$, i.e. $b = 0$. So the eigenvalue is real, and so we may also take real eigenvectors.

<u>part 2:</u> Eigenspaces for different eigenvalues are orthogonal (class notes November 30, text section 7.1).

<u>proof:</u>
Let

$$A\,\underline{\boldsymbol{v}} = \lambda_1 \underline{\boldsymbol{v}} \qquad A\,\underline{\boldsymbol{w}} = \lambda_2\,\underline{\boldsymbol{w}}$$

with $\lambda_1 \neq \lambda_2$. Then we need to verify that $\underline{\boldsymbol{v}} \perp \underline{\boldsymbol{w}}$.

Proceding as in <u>part 1</u>, consider the scalar

$$\underline{\boldsymbol{v}}^T A\,\underline{\boldsymbol{w}}.$$

On one hand

$$\underline{\boldsymbol{v}}^T (A\,\underline{\boldsymbol{w}}) = \underline{\boldsymbol{v}}^T\left(\lambda_2\,\underline{\boldsymbol{w}}\right) = \lambda_2\,\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}}.$$

On the other hand

$$\left(\underline{\boldsymbol{v}}^T A\right)\underline{\boldsymbol{w}} = \left(\underline{\boldsymbol{v}}^T A^T\right)\underline{\boldsymbol{w}} = (A\,\underline{\boldsymbol{v}})^T\underline{\boldsymbol{w}} = \lambda_1\,\underline{\boldsymbol{v}}^T\,\underline{\boldsymbol{w}} = \lambda_1\,\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}}\ .$$

So

$$\lambda_2\,\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}} = \lambda_1\,\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}}.$$

So as long as $\lambda_1 \neq \lambda_2$ we deduce that the eigenvectors $\underline{\boldsymbol{v}}$, $\underline{\boldsymbol{w}}$ satsify $\underline{\boldsymbol{v}} \cdot \underline{\boldsymbol{w}} = 0$!

part 3: (not in text) We prove the spectral theorem by induction on the size of A. The case in which A is a 1×1 matrix is true, since the number 1 is an eigenvector, and its eigenvalue is whatever the scalar value of A is. Now, assume the spectral theorem holds for all symmetric matrices of size $(n-1) \times (n-1)$ and lower. We will prove it for $n \times n$ matrices, based on the inductive hypothesis. So, let A be $n \times n$ and symmetric, with $n \geq 2$. Compute the characteristic polynomial and pick one of its real roots λ_1. Find the corresponding eigenspace $E_{\lambda=\lambda_1}$ and use Gram-Schmidt to find an orthonormal eigenbasis. If $E_{\lambda=\lambda_1}$ is all of $\mathbb{R}^n$ then $A\,\underline{\boldsymbol{v}} = \lambda_1\,\underline{\boldsymbol{v}}$ for all vectors $\underline{\boldsymbol{v}}$ and the standard basis is an orthonormal eigenbasis.

Otherwise, let $\{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \underline{\boldsymbol{u}}_k\}$ be an orthonormal basis for $E_{\lambda=\lambda_1}$ and let $\{\underline{\boldsymbol{u}}_{k+1}, \ldots \underline{\boldsymbol{u}}_n\}$ be an orthonormal basis for $W = \left(E_{\lambda=\lambda_1}\right)^{\perp}$. Combining these bases, we have an orthonormal basis for $\mathbb{R}^n$,

$$\beta = \{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \underline{\boldsymbol{u}}_k, \underline{\boldsymbol{u}}_{k+1}, \ldots \underline{\boldsymbol{u}}_n\},$$

although the last $n-k$ vectors are not yet eigenvectors. We will use the inductive hypothesis to take care of that, as follows.

First notice that the linear transformation

$$T\,(\underline{\mathbf{x}}) = A\,\underline{\mathbf{x}}$$

preserves the orthogonal complement W to $E_{\lambda=\lambda_1}$ because A is symmetric. This is because for $\underline{\mathbf{w}} \in W$, and $\underline{\mathbf{v}} \in E_{\lambda=\lambda_1}$,

$$A\,\underline{\mathbf{w}} \cdot \underline{\mathbf{v}} = (A\,\underline{\mathbf{w}})^T\underline{\mathbf{v}} = \underline{\mathbf{w}}^T A\,\underline{\mathbf{v}} = \underline{\mathbf{w}}^T\lambda_1\,\underline{\mathbf{v}} = \lambda_1\underline{\mathbf{w}} \cdot \underline{\mathbf{v}} = 0.$$

In other words,

$$\underline{\mathbf{w}} \in \left(E_{\lambda=\lambda_1}\right)^{\perp} \Rightarrow A\,\underline{\mathbf{w}} \in \left(E_{\lambda=\lambda_1}\right)^{\perp}.$$

Now apply this information to the matrix for T with respect to β: Because T preserves the orthogonal complement of $E_{\lambda=\lambda_1}$, and because the first k vectors of β are eigenvectors,

$$[T]_\beta = \left[\ [A\,\underline{\boldsymbol{u}}_1]_\beta\ [A\,\underline{\boldsymbol{u}}_2]_\beta\ \cdots\ [A\,\underline{\boldsymbol{u}}_n]_\beta\ \right]$$

$$= \begin{bmatrix} \lambda_1 I_{k \times k} & [0]_{k \times n-k} \\ [0]_{(n-k)\times k} & B_{(n-k)\times(n-k)} \end{bmatrix}.$$

In other words, the first k columns are just like those of an $n \times n$ matrix which is λ_1 times the identity. And the last $n - k$ columns have entries which are zero in the first k rows. On the other hand, for the orthogonal change of basis matrix P with columns $\{\underline{\boldsymbol{u}}_1, \underline{\boldsymbol{u}}_2, \underline{\boldsymbol{u}}_k, \underline{\boldsymbol{u}}_{k+1}, \cdots \underline{\boldsymbol{u}}_n\}$

$$[T]_\beta = P^{-1} A\,P = P^T A\,P.$$

Because $P^T A\,P$ is symmetric (take its transpose to check!), we deduce that the submatrix $B_{(n-k)\times(n-k)}$ is also symmetric. *By the inductive hypothesis* this smaller matrix has an orthonormal eigenbasis. These eigenbasis vectors for the matrix B are the β-coordinates of an orthonormal eigenbasis for W. (You can check this.) Combining this eigenbasis for $W = \left(E_{\lambda=\lambda_1}\right)^\perp$ with the original eigenbasis for $E_{\lambda=\lambda_1}$ yields an orthonormal eigenbasis for $\mathbb{R}^n$.

QED.