Math 204. Final Exam Review.

April 27, 2010

Here is a summary of the facts. Of course you need to know how to use these facts, and where they come from. That's what the class and book are for.

1. $\mathbb{R}^n$ denotes the standard n-dimensional vector space consisting of elements of the form $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, where $x_1, \ldots, x_n \in \mathbb{R}$. The standard basis for this space is given by the vectors $\vec{e}_i$, where $\vec{e}_i$ has a one in the ith entry and is zero otherwise.

2. The set of $m \times n$ matrices is denoted $\mathbb{R}^{m\times n}$. Given a matrix $A \in \mathbb{R}^{m\times n}$, there are two natural ways to represent it. We can either write it in terms of its columns as $A = [\vec{C}_1|\cdots|\vec{C}_n]$ or in terms of its rows $A = \begin{bmatrix} \vec{R}_1 \\ \vdots \\ \vec{R}_m \end{bmatrix}$. The product of A and $\vec{x} \in \mathbb{R}^n$ is given by

$$A\vec{x} = x_1\vec{C}_1 + \cdots + x_n\vec{C}_n = \begin{bmatrix} \vec{R}_1 \cdot \vec{x} \\ \vdots \\ \vec{R}_m \cdot \vec{x} \end{bmatrix}.$$

3. There are three elementary row operations
   (a) $R_i \leftrightarrow R_j$
   (b) $R_i \leftarrow R_i + cR_j$, where $i \neq j$ and $c \in \mathbb{R}$.
   (c) $R_i \leftarrow kR_i$, where $k \neq 0$.

4. Given a matrix $A \in \mathbb{R}^{m\times n}$ there is a sequence of elementary row operations that brings A to row-reduced echelon form. A matrix $A \in \mathbb{R}^{m\times n}$ is in row-reduced echelon form if the following conditions are true:
   (a) All zero rows are at the bottom of the matrix.
   (b) The first nonzero entry in each nonzero row is a 1 (called the leading 1).
   (c) If a column contains a leading 1, then all the other entries in that column are 0.
   (d) If the ith and jth rows contain a leading 1 in columns k and l and if $i < j$, then $k < l$.

   The row-reduced echelon form is unique.

5. Every row operation corresponds to an invertible $m \times m$ matrix E such that EA is the matrix obtained from A by applying the row operation that corresponds to E.

6. A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a linear transformation if and only if $T(c\vec{x} + \vec{y}) = cT(\vec{x}) + T(\vec{y})$ for every $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

7. There is a one-to-one correspondence between linear transformations $T : \mathbb{R}^n \to \mathbb{R}^m$ and matrices $A \in \mathbb{R}^{m\times n}$, given by $T(\vec{x}) = A\vec{x}$.

8. If S, T are linear transformations from $\mathbb{R}^n \to \mathbb{R}^m$, then $(S + T)$ is defined by $(S + T)\vec{x} = S\vec{x} + T\vec{x}$. We also define cT by $(cT)(\vec{x}) = cT(\vec{x})$. If $S : \mathbb{R}^p \to \mathbb{R}^m$ and $T : \mathbb{R}^n \to \mathbb{R}^p$, then we define $ST = S \circ T : \mathbb{R}^n \to \mathbb{R}^m$ by $(S \circ T)(\vec{x}) = S(T(\vec{x}))$.

   These operations correspond to matrix addition, multiplication by a scalar, and matrix multiplication. Matrix multiplication is not commutative, i.e., there are examples of matrices A, B such that $AB \neq BA$.

9. Example of matrices that we have seen in class and in the homework:

   (a) Projections onto lines and planes
   (b) Rotations
   (c) Reflections
   (d) Shears
   (e) Scaling
   (f) Triangular matrices
   (g) Diagonal matrices

10. A non-empty subset $W \subset \mathbb{R}^n$ is called a subspace if and only if it is closed under linear combinations, i.e., if $\vec{x}, \vec{y} \in W$ and $c \in \mathbb{R}$, then the vector $c\vec{x} + \vec{y} \in W$.

    Examples of subspaces include

    (a) The null-space or kernel of a matrix (linear transformation) $A \in \mathbb{R}^{m\times n}$. This is the set of all vectors in $\mathbb{R}^n$ such that $A\vec{x} = 0$, i.e., $\text{kernel}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = 0\}$. Note that the null space is a subspace of $\mathbb{R}^n$.
    (b) The range or image of A is defined by $\text{range}(A) = \{A\vec{x} : x \in \mathbb{R}^n\}$. Note that $A\vec{x} \in \mathbb{R}^m$ and so the range is a subspace of $\mathbb{R}^m$.
    (c) Given any collection of vectors $\vec{v}_1, \dots, \vec{v}_m \in \mathbb{R}^m$ we can build a subspace by taking all possible linear combinations of these vectors. This is called the span of the vectors and is defined by $\text{span}\{\vec{v}_1, \dots, \vec{v}_m\} = \{c_1\vec{v}_1 + \cdots + c_m\vec{v}_m : c_1, \dots, c_m \in \mathbb{R}\}$.

11. The range of a matrix is the span of its columns. You can find a basis for this space by computing $\text{rref}(A)$. The columns of A that have a leading one in $\text{rref}(A)$ constitute a basis for the range of A. A basis for the null space can be found by solving $A\vec{x} = 0$.

12. We have the following equation

$$\begin{aligned} \text{number of free variables} + \text{number of leading ones} &= n \\ \text{dimension of the null space} + \text{dimension of the range} &= n \\ \text{nullity}(A) + \text{rank}(A) &= n \end{aligned}$$

13. A matrix $A \in \mathbb{R}^{n\times n}$ is invertible if and only if there is a matrix B such that $AB = BA = I_n$. This is equivalent to each of the following statements

    (a) There exists a matrix B such that $AB = I_n$
    (b) There exists a matrix B such that $BA = I_n$.
    (c) $\text{rank}(A) = n$
    (d) $\text{nullity}(A) = 0$.
    (e) $\text{rref}(A) = I_n$
    (f) $\text{range}(A) = \mathbb{R}^n$
    (g) $\text{kernel}(A) = \{\vec{0}\}$.
    (h) The system $A\vec{x} = \vec{b}$ has a unique solution for every choice of $\vec{b}$.

14. Let $\{\vec{v}_1, \dots, \vec{v}_m\}$ be a collection of vectors in $\mathbb{R}^n$. The span of $\{\vec{v}_1, \dots, \vec{v}_m\}$ is the set of all linear combinations of the vectors $\{\vec{v}_1, \dots, \vec{v}_m\}$. We denote the span by $\text{span}\{\vec{v}_1, \dots, \vec{v}_m\}$. This set is called

    (a) linearly independent if $c_1\vec{v}_1 + \cdots + \vec{v}_m = \vec{0}$ implies that $c_1 = \cdots = c_m = 0$.
    (b) linearly dependent if it is not linearly independent.
    (c) spanning for a subspace V if $V = \text{span}\{\vec{v}_1, \dots, \vec{v}_m\}$.

15. Suppose that $W \subseteq \mathbb{R}^n$ is a subspace. If $\{v_1, \dots, v_p\}$ is a linearly independent set in W, and $\{w_1, \dots, w_m\}$ is a spanning set for W, then $p \le m$. A basis for W is linearly independent, spanning set. The dimension of W is the number of elements in any basis of W.

16. The orthogonal complement of a set $S \subseteq \mathbb{R}^n$, denoted $S^\perp$, is defined by $S^\perp = \{w \in \mathbb{R}^n : w \cdot x = 0 \text{ for all } x \in S\}$.

17. If V, W are subspaces of $\mathbb{R}^n$, then $V + W := \{\vec{v} + \vec{w} : \vec{v} \in V, \vec{w} \in W\}$. $V + W$ is a subspace and $\dim(V + W) + \dim(V \cap W) = \dim(V) + \dim(W)$.

18. If W is a subspace of $\mathbb{R}^n$, then $W \cap W^\perp = \{\vec{0}\}$ and $W + W^\perp = \mathbb{R}^n$.

19. A set of vectors $\{\vec{u}_1, \dots, \vec{u}_p\}$ is called orthonormal if $\|\vec{u}_i\| = 1$ and $\vec{u}_i \cdot \vec{u}_j = 0$ whenever $i \neq j$. The set is called an orthonormal basis for W if it is orthonormal and spans W. An orthonormal set of vectors is linearly independent.

20. If $\{\vec{u}_1, \dots, \vec{u}_m\}$ is an orthonormal set, then the projection onto $W = \text{span}\{\vec{u}_1, \dots, \vec{u}_m\}$ is given by
$$\text{proj}_W(\vec{x}) = (\vec{x} \cdot \vec{u}_1)\vec{u}_1 + \cdots + (\vec{x} \cdot \vec{u}_m)\vec{u}_m$$

21. Given a linearly independent set $\{\vec{v}_1, \dots, \vec{v}_m\}$ the Gram-Schmidt process produces an orthonormal set $\{\vec{u}_1, \dots, \vec{u}_m\}$ such that $\text{span}\{\vec{u}_1, \dots, \vec{u}_k\} = \text{span}\{\vec{v}_1, \dots, \vec{v}_k\}$ for $1 \le k \le m$.

    To carry out the Gram-Schmidt process we set $V_0 = \{\vec{0}\}$ and $V_k = \text{span}\{\vec{u}_1, \dots, \vec{u}_k\}$ and let $\vec{u}_k = \alpha_k(\vec{v}_k - \text{proj}_{V_{k-1}}(\vec{v}_k))$, the constant α_k is chosen so that $\|\vec{u}_k\| = 1$.

22. If $A \in \mathbb{R}^{m\times n}$, and $A = [a_{i,j}]$, then $A^T = [a_{j,i}]$, is called the transpose of A. A matrix is called symmetric if $A = A^T$. Basic properties $(AB)^T = B^T A^T$, $(A^T)^T = A$, $(A^{-1})^T = (A^T)^{-1}$.

23. A square $n \times n$ matrix A is called an orthogonal matrix if the columns of A form an orthonormal basis for $\mathbb{R}^n$. An orthogonal matrix preserves dot products and lengths. The inverse of an orthogonal matrix is A^T. The product of orthogonal matrices is orthogonal, as is the inverse.

24. We have $\text{range}(A)^\perp = \ker(A^T)$ and $\ker(A^T A) = \ker(A)$.

25. Given a matrix $A \in \mathbb{R}^{m\times n}$ and a vector $\vec{b}$, the least-squares problem is to find a vector $\vec{x}^*$ that minimizes $\|\vec{b} - Ax\|$. The set of solutions to this problem is given by the consistent system $A^T A\vec{x} = A^T\vec{b}$. This is called the normal equation. If $\ker(A) = \{\vec{0}\}$, then $A^T A$ is invertible, and there is a unique solution $\vec{x}^* = (A^T A)^{-1}A^T\vec{b}$.

    Otherwise we define the minimal least-squares solution to be the vector $\vec{x}^{**}$ of minimal length that solves $A^T A\vec{x} = A^T\vec{b}$. This is the unique vector in $\ker(A)^\perp$ that solves the normal equations. It is given by $\vec{x}^{**} = \text{proj}_{\ker(A)^\perp}(\vec{x}^*)$.

26. A real vector space (linear space) V is a set together with two operations called addition and scalar multiplication. The operations must satisfy the following rules for $x, y, z \in V$ and $c, d \in \mathbb{R}$.:

    (a) $x + y = y + x$ commutative
    (b) $(x + y) + z = x + (y + z)$ associative
    (c) there is vector n such that $x + n = x$ for all $x \in V$ neutral element
    (d) For every vector x there is a vector y such that $x + y = n$ inverse
    (e) $(c + d)x = cx + dx$
    (f) $c(x + y) = cx + cy$
    (g) $(cd)x = c(dx)$
    (h) $1x = x$

27. Examples of vector spaces include $\mathbb{R}^n$, subspaces of $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$, symmetric matrices, upper-triangular matrices, diagonal matrices, polynomials, polynomials of degree at most n, continuous functions on the set $[0, 1]$, differentiable functions on $\mathbb{R}$.

28. Given a matrix $A \in \mathbb{R}^n$ we write $A = [C_1|\cdots|C_n]$ where C_i is the ith column of A. A function $f : \mathbb{R}^{n\times n} \to \mathbb{R}$ is called multilinear if, when $n - 1$ of the columns are fixed, the function f is linear is linear in the remaining column. It is called alternating if the $f(A) = -f(B)$, whenever B is obtained from A by interchanging two different columns. The only multilinear, alternating function on $\mathbb{R}^{n\times n}$ such that $f(I_n) = 1$ is called the determinant det. Here are some properties of the determinant:

    (a) $\det(kA) = k^n \det(A)$.
    (b) $\det(AB) = \det(A)\det(B)$.
    (c) $\det(A^T) = \det(A)$.
    (d) $\det(A) = 0$ if and only if A is singular.
    (e) $\det(A^{-1}) = \dfrac{1}{\det(A)}$ whenever A is invertible.

(f) The determinant of an upper-triangular matrix is the product of its diagonal entries.

If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$, the determinant is given by

$$\det(A) = ad - bc.$$

We can define the determinant recursively (or inductively) by the cofactor rule

$$\det\left(\begin{bmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{bmatrix}\right) = \sum_{k=1}^{n} (-1)^{(k-1)} a_{1,k} \det(A_{1,k}),$$

where $A_{1,k}$ is the matrix obtained from A by "deleting" the first row and kth column.

29. The trace of a matrix $A \in \mathbb{R}^n$ is defined by $\operatorname{trace}(A) = a_{1,1} + \cdots + a_{n,n}$. The trace is a linear transformation from $\mathbb{R}^{n\times n} \to \mathbb{R}$. In addition, the trace satisfies $\operatorname{trace}(AB) = \operatorname{trace}(BA)$.

30. Let A be an $n \times n$ matrix. A scalar λ is called an eigenvalue of A if there exists a nonzero vector x such that $Ax = \lambda x$.

    The charactersitic polynomial of A is defined by $p_A(t) = \det(A - tI)$. $p_A(t)$ is an nth degree polynomial.

    The following conditions are equivalent

    (a) λ is an eigenvalue of A,
    (b) $\ker(A - \lambda I) \neq \{0\}$,
    (c) $A - \lambda I$ is singular
    (d) $\det(A - \lambda I) = 0$
    (e) λ is a root of the characteristic polynomial $p_A(t)$.

    The eigenspace $E_\lambda = \ker(A - \lambda I)$. If $\lambda \neq \mu$, then $E_\lambda \cap E_\mu = \{0\}$.

31. For an $n \times n$ matrix A we have $\det(A) = \lambda_1 \cdots \lambda_n$, and $\operatorname{trace}(A) = \lambda_1 + \cdots + \lambda_n$, where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A counting multiplicity.

    In general, $p_A(t) = (-1)^n t^n + (-1)^{n-1} \operatorname{trace}(A) + \cdots + \det(A)$.

    This is particularly useful in the 2×2 case, since we can write $p_A(t) = t^2 - \operatorname{trace}(A)t + \det(A)$.

32. The fundamental theorem of algebra states that every nth degree polynomial has a factorization of the form $p(t) = (t - \lambda_1)^{a_1} \cdots (t - \lambda_l)^{a_l} q(t)$, where the numbers $\lambda_1, \ldots, \lambda_l$ are distinct and the polynomial q has no real roots. The numbers $a_1, \ldots, a_l$ are called the algebraic multiplicities of the eigenvalues $\lambda_1, \ldots, \lambda_l$. The geometric multiplicity of an eigenvalue λ_i is defined to be the dimension of the corresponding eigenspace E_{λ_i}. For the purposes of these notes we will denote the geometric multiplicity by g_i.

    It is always the case that $g_i \leq a_i$.

33. If $\mathcal{B} = \{v_1, \ldots, v_n\}$ is a basis of $\mathbb{R}^n$, then the coordinates of a vector x in the basis $\mathcal{B}$ is the vector $[x]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$, where $x = c_1 v_1 + \cdots c_n v_n$.

    If we let $S = [v_1|\cdots|v_n]$, then S is invertible and is called the change of basis matrix. If $A \in \mathbb{R}^n$, then there is aunique matrix B such that $[Ax]_{\mathcal{B}} = B[x]_{\mathcal{B}}$. The matrices A and B are related by $S^{-1}AS = B$.

    We say that A is similar to B if there is an invertible matrix S such that $AS = SB$. Similarity of matrices is reflexive (A is similar to A), symmetric (If A is similar to B, then B is similar to A), and transitive (If A is similar to B, an B is similar to C, then A is similar to C).

    Similar matrices represent the same linear transformation in different bases.

    If A and B are similar, then they have the same characteristic polynomial, and eigenvalues. In addition, the algebraic multiplicities and geometric multiplicities are preserved. These facts follow from the observation that $A - \lambda I$ is similar to $B - \lambda I$.

34. An eigenbasis for A is a basis of $\mathbb{R}^n$, in which each vector is an eigenvector of A. If $\mathcal{B} = \{v_1, \ldots, v_n\}$ is an eigenbasis for A, then the matrix A is diagonal in the basis $\mathcal{B}$. The diagonal entries of this matrix are the eigenvalues of the matrix A. A matrix is called diagonalizable if there is an eigenbasis for A.

    A matrix $A \in \mathbb{R}^{n\times n}$ is diagonalizable if and only if the algebraic multiplicities of the eigenvalues of A sum to n, i.e., $n = a_1 + \cdots + a_l$ and for each eigenvalue λ the geometric multiplicty of λ is equal to the algebraic multiplicity of λ.

35. The spectral theorem tells us that a symmetric matrix A is diagonalizable. In addition, if $\lambda \neq \mu$, then $E_\lambda \perp E_\mu$. Therefore, A has an orthonormal eigenbasis and the change of basis matrix can be chosen to be orthogonal. Therefore, there is an orthogonal matrix U such that $U^T AU = D$, where the diagonal entries of D are the eigenvalues of A