Math 210 F **Spring 2016**
Study Sheet for the Final **Drew Armstrong**

After spending all semester emphasizing conceptual understanding over memorization, here is a list of formulas you can memorize for the final exam. Sorry there are no pictures; those take a long time to make on the computer.

Points: Let $\mathbb{R}^n$ denote the set of $n \times 1$ matrices of real numbers:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

We call these "points" in n-dimensional Cartesian space.

Vectors: We will also think of a point $\vec{x}$ in $\mathbb{R}^n$ as a directed line segment (a "vector") with its tail at the origin $\vec{0}$ and its head at the point $\vec{x}$. This idea is subtle because we are allowed to pick up the arrow and move it as long as we don't change its length or direction.

Parallolgram Law: Consider two points $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$. The points $\vec{0}, \vec{x}, \vec{y}$ form three vertices of a 2D parallelogram living in $\mathbb{R}^n$. The fourth vertex of the parallogram is

$$\vec{x} + \vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Subtraction of Vectors: Consider two points $\vec{x}, \vec{y}$ in $\mathbb{R}^n$. The vector with tail at $\vec{x}$ and head at $\vec{y}$ is represented by the point

$$\vec{y} - \vec{x} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} - \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \\ \vdots \\ y_n - x_n \end{pmatrix}.$$

Vector Arithmetic: Consider vectors $\vec{x}, \vec{y}, \vec{z}$ in $\mathbb{R}^n$ and numbers a, b in $\mathbb{R}$. Then we have

- $\vec{x} + \vec{y} = \vec{y} + \vec{x}$,
- $\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}$,
- $a(\vec{x} + \vec{y}) = a\vec{x} + a\vec{y}$,
- $(a + b)\vec{x} = a\vec{x} + b\vec{x}$.

Dot Product: Given vectors $\vec{x}, \vec{y}$ in $\mathbb{R}^n$, we define their dot product as the number

$$\vec{x} \bullet \vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \bullet \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

More Vector Arithmetic: For all vectors $\vec{x}, \vec{y}, \vec{z}$ in $\mathbb{R}^n$ and numbers a in $\mathbb{R}$ we have

– $\vec{x} \bullet \vec{y} = \vec{y} \bullet \vec{x}$,
– $\vec{x} \bullet (\vec{y} + a\vec{z}) = \vec{x} \bullet \vec{y} + a\vec{x} \bullet \vec{z}$.

Pythagorean Theorem: Given a vector $\vec{x}$ in $\mathbb{R}^n$ its "length" $\|\vec{x}\|$ is the non-negative number defined by

$$\|\vec{x}\|^2 = \vec{x} \bullet \vec{x} = x_1^2 + x_2^2 + \cdots x_n^2.$$

Law of Cosines: Consider two vectors $\vec{x}, \vec{y}$ in $\mathbb{R}^n$. These vectors together with their difference $\vec{y} - \vec{x}$ form the three sides of a 2D triangle in $\mathbb{R}^n$. By applying the formulas above we get

$$\|\vec{y} - \vec{x}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2(\vec{x} \bullet \vec{y}).$$

On the other hand, the classical Law of Cosines for triangles tells us that

$$\|\vec{y} - \vec{x}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2\|\vec{x}\|\|\vec{y}\| \cos\theta,$$

where θ is the angle between the vectors $\vec{x}$ and $\vec{y}$. Then comparing the two equations gives

$$\vec{x} \bullet \vec{y} = \|\vec{x}\|\|\vec{y}\| \cos\theta.$$

In particular, this tells us that $\vec{x} \perp \vec{y}$ if and only if $\vec{x} \bullet \vec{y} = 0$.

Lines in $\mathbb{R}^2$: A line in the plane can be written in parametric form as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} + t \begin{pmatrix} u \\ v \end{pmatrix}.$$

This is the line containing the point (x_0, y_0) and parallel to the vector (u, v). Or it can be expressed by an equation

$$ax + by = c$$

where (a, b) is some vector perpendicular ("normal") to the line. This line contains the origin $(0, 0)$ if and only if $c = 0$. In general, the line has minimum distance $c/\sqrt{a^2 + b^2}$ from the origin.

Planes in $\mathbb{R}^3$: A plane in 3-dimensional space can be written in parametric form as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + s \begin{pmatrix} u_1 \\ v_1 \\ w_1 \end{pmatrix} + t \begin{pmatrix} u_2 \\ v_2 \\ w_2 \end{pmatrix}.$$

This is the plane containing the point (x_0, y_0, z_0) and spanned by the vectors (u_1, v_1, w_1) and (u_2, v_2, w_2). Or it can be expressed by an equation

$$ax + by + cz = d$$

where (a, b, c) is some vector perpendicular ("normal") to the plane. This plane contains the origin $(0, 0, 0)$ if and only if $d = 0$. In general, the plane has minimum distance $d/\sqrt{a^2 + b^2 + c^2}$ from the origin.

Lines in $\mathbb{R}^3$: A line in 3-dimensional space can be written in parametric form as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t \begin{pmatrix} u \\ v \\ w \end{pmatrix}.$$

This is the line containing the point (x_0, y_0, z_0) and parallel to the vector (u, v, w). However, a line in 3D can **not** be defined by a single equation. It **can** be defined as the solution of a system of two linear equations in three unknowns. Geometrically, this expresses the line as an intersection of two planes.

Systems of Linear Equations: A system of m linear equations in n unknowns has the following form:

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ & & & & & & & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{cases}$$

Alternatively, we can write it as a matrix equation:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

And we usually shorten it to this:

$$A\vec{x} = \vec{b}.$$

The set of solutions to this system form a flat d-dimensional shape (called a "d-plane") living in $\mathbb{R}^n$. Most likely we will have $d = n - m = \#$ variables $- \#$ equations.

Gaussian Elimination: A system of linear equations $A\vec{x} = \vec{b}$ can be solved by putting the system in Reduced Row Echelon Form (RREF) by Gaussian elimination. Each non-pivot column in the RREF leads to a free variable, so if there are d non-pivot columns in the RREF the solution will be a d-plane. Each non-pivot column also tells us an explicit non-trivial relation among the columns of A. We call d the "nullity" of A, and write $d = \text{null}(A)$.

Fundamental Theorem I: The set of vectors of the form $A\vec{x}$ is called the column space of the matrix A, because it consists of all linear combinations of the columns. If A has shape $m \times n$ then the column space is a subspace of $\mathbb{R}^m$. The dimension of the column space is called the "rank" of A, written $\text{rank}(A)$. It is equal to the number of pivot columns in the RREF. Since the total number of columns in the RREF is n, we obtain

$$\text{rank}(A) + \text{null}(A) = n.$$

Fundamental Theorem II: Let A have shape $m \times n$. Many times have I told you that the equation $A^T\vec{e} = \vec{0}$ means that the vector $\vec{e}$ is perpendicular to all of the columns of A. In other words, the nullspace of A^T is the "orthogonal complement" to the column space of A. It follows from this that their dimensions add to m, i.e., $\text{null}(A^T) + \text{rank}(A) = m$. Combining this with the Fundamental Theorem above gives the following surprising equation:

$$\text{rank}(A^T) = \text{rank}(A).$$

In other words, the row space and the column space of A have the same dimension!

Matrix Multiplication: Let A have shape $\ell \times m$ and let B have shape $m \times n$. Then the matrix AB exists and has shape $\ell \times n$. It is defined by requiring that the following equation holds for all $\vec{x}$ in $\mathbb{R}^n$:

$$(AB)\vec{x} = A(B\vec{x}).$$

However, if we want to actually **compute** the matrix AB we use the following rules:

- ((i,j)-th entry of AB) = (i-th row of A)(j-th column of B)
- (i-th row of AB) = (i-th row of A) B
- (j-th column of AB) = A (j-th column of B).

Inverse Matrices: Let A have shape $m \times n$. We say that B is an inverse matrix of A if $AB = I_m$ and $BA = I_n$. But this is impossible unless $m = n$. (Reason: The matrix B has shape $n \times m$. If $n < m$ then RREF(B) has a non-pivot column so B has a non-trivial column relation. But then since A(j-th column B)=(j-th column I_m) we conclude that I_m has a non-trivial column relation, which is impossible. If $m < n$ then A has a non-trivial column relation and then the equation $BA = I_n$ tells us that I_n has a non-trivial column relation, which is impossible.) If A has shape $n \times n$ then it **might** have an inverse. To compute the inverse we do this trick:

$$(A|I) \xrightarrow{\text{RREF}} \left(I|A^{-1}\right)$$

If the trick doesn't work (because A had some non-trivial row relation or column relation) then we conclude that A has no inverse.

Uniqueness of Inverses: Suppose that we have $AB = I$ and $CA = I$. It follows that

$$C = CI = C(AB) = (CA)B = IB = B.$$

Hence if A has an inverse matrix, this matrix is unique. We give it the special name A^{-1}.

Matrix Arithmetic: Consider matrices A, B, C and numbers x, y. The following formulas hold as long as the respective matrices exist:

- $A + B = B + A$
- $A + (B + C) = (A + B) + C$
- $A(BC) = (AB)C$
- $(x + y)A = xA + yA$
- $x(AB) = (xA)B = A(xB)$
- $A(B + xC) = AB + xAC$
- $(A + xB)C = AC + xBC$
- $(A + B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$.

WARNING: The following two formulas are NOT generally true:

- $AB = BA$.
- $(A + B)^{-1} = A^{-1} + B^{-1}$.

Solutions of a Linear System are Flat: Suppose we have two solutions of a linear system: $A\vec{x} = \vec{b}$ and $A\vec{y} = \vec{b}$. Then for any number t we have

$$A(t\vec{x} + (1 - t)\vec{y}) = tA\vec{x} + (1 - t)A\vec{y} = t\vec{b} + (1 - t)\vec{b} = \vec{b}.$$

This implies that every point of the line $t\vec{x} + (1 - t)\vec{y}$ is also a solution. This is what I mean when I say that the solutions of a linear system form a d-plane. Geometrically: If m hyperplanes in n-dimensional space meet at two points $\vec{x}, \vec{y}$ then they also meet at the whole line $t\vec{x} + (1 - t)\vec{y}$.

Orthogonal Projection: Let A be an $m \times n$ matrix such that the inverse $(A^T A)^{-1}$ exists. Then the $m \times m$ matrix

$$P = A(A^T A)^{-1} A^T$$

satisfies the properties $P^2 = P$ and $P^T = P$. Geometrically, this matrix projects any point orthogonally onto the column space of A. Special case: If $A = \vec{a}$ is a column vector then we have

$$P = \vec{a}(\vec{a}^T\vec{a})^{-1}\vec{a}^T = \frac{1}{\vec{a}^T\vec{a}}\vec{a}\vec{a}^T = \frac{1}{\|\vec{a}\|^2}\vec{a}\vec{a}^T.$$

This is the matrix that projects onto the line $t\vec{a}$. If Q is the matrix that projects onto the nullspace of A^T (which consists of all vectors perpendicular to the column space of A) then we have

$$P + Q = I_m.$$

Least Squares Regression: Suppose that the linear system $A\vec{x} = \vec{b}$ has no solution. This means that the point $\vec{b}$ is not in the column space of A. Gauss' idea was to project the point $\vec{b}$ orthogonally into the column space of A to get $P\vec{b}$ and then to solve the equation $A\hat{x} = P\vec{b}$ instead. This new "normal equation" is usually written as

$$A^T A\hat{x} = A^T\vec{b}.$$

The most common application of this equation is to find the line that is a best fit for a given set of data points.

Eigenvectors and Eigenvalues: Let A be an $n \times n$ matrix. We say that a nonzero vector $\vec{x} \neq \vec{0}$ is an eigenvector of A if there exists a number λ such that

$$A\vec{x} = \lambda\vec{x}.$$

In this case we say that $\vec{x}$ is an eigenvector with eigenvalue λ, or a λ-eigenvector. This equation can be rewritten as

$$(A - \lambda I_n)\vec{x} = \vec{0}.$$

Note that this equation has a non-zero solution (i.e., λ is an eigenvalue) precisely when the matrix $(A - \lambda I_n)$ is **not invertible**. In the case of a 2×2 matrix we can express this as the characteristic equation

$$0 = \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} = \lambda^2 - (a + d)\lambda + (ad - bc).$$

Spectral Analysis: Suppose you have a linear recurrence relation defined by $\vec{v}_{n+1} = A\vec{v}_n$. If $\vec{v}_0$ is the initial condition then the n-th state vector is given by

$$\vec{v}_n = A^n\vec{v}_0.$$

To solve this equation we first find the eigenvalues of A via the characteristic equation and then we find some corresponding eigenvectors. Suppose we find

$$A\vec{x} = \lambda\vec{x} \qquad \text{and} \qquad A\vec{y} = \mu\vec{y}.$$

Then we try to express our initial condition in terms of eigenvectors: $\vec{v}_0 = a\vec{x} + b\vec{y}$. If we're successful (i.e., if the matrix A has enough eigenvectors) then we can use this to obtain a "closed form" solution to the recurrence:

$$\begin{aligned} \vec{v}_n = A^n\vec{v}_0 &= A^n(a\vec{x} + b\vec{y}) \\ &= aA^n\vec{x} + bA^n\vec{y} \\ &= a\lambda^n\vec{x} + b\mu^n\vec{y}. \end{aligned}$$