# Math 314

## Topics for second exam

Technically, everything covered by the first exam **plus**

**Chapter 2** §6 Determinants

(Square) matrices come in two flavors: invertible (all $Ax = b$ have a solution) and non-invertible ($Ax = \mathbf{0}$ has a non-trivial solution). It is an amazing fact that one number identifies this difference; the determinant of $A$.

For 2×2 matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, this number is det($A$)=$ad - bc$; if $\neq 0$, $A$ is invertible, if =0, $A$ is non-invertible (=singular).

For larger matrices, there is a similar (but more complicated formula):

$A$= $n \times n$ matrix, $M_{ij}(A)$ = matrix obtained by removing $i$th row and $j$th column of $A$.

$\det(A) = \Sigma_{i=1}^{n}(-1)^{i+1}a_{i1}\det(M_{i1}(A))$

(this is called expanding along the first column)

Amazing properties:

If $A$ is upper triangular, then $\det(A)$ = product of the entries on the diagonal

If you multiply a row of $A$ by $c$ to get $B$, then $\det(B) = c\det(A)$

If you add a mult of one row of $A$ to another to get $B$, then $\det(B) = \det(A)$

If you switch a pair of rows of $A$ to get $B$, then $\det(B) = -\det(A)$

In other words, we can understand exactly how each elementary row operation affects the determinant. In part,

$A$ is invertible iff $\det(A) \neq 0$; and in fact, we can **use** row operations to calculate $\det(A)$ (since the RREF of a matrix is upper triangular).

More interesting facts:

$\det(AB) = \det(A)\det(B)$ ; $\det(A^T) = \det(A)$

We can expand along other columns than the first:

$\det(A) = \Sigma_{i=1}^{n}(-1)^{i+j}a_{ij}\det(M_{ij}(A))$

(expanding along $j$th column)

And since $\det(A^T) = \det(A)$, we could expand along **rows**, as well....

A formula for the inverse of a matrix:

If we define $A_c$ to be the matrix whose $(i,j)$th entry is $(-1)^{i+j}\det(M_{ij}(A))$, then $A_c^T A = (\det A)I$ ($A_c^T$ is called the *adjoint* of $A$). So if $\det(A) \neq 0$, then we can write the inverse of $A$ as

$A^{-1} = \dfrac{1}{\det(A)}A_c^T$ (This is very handy for 2×2 matrices...)

The same approach allows us to write an explicit formula for the solution to $Ax = b$, when $A$ is invertible:

If we write $B_i = A$ with its $i$th column replaced by $b$, then the (unique) solution to $Ax = b$ has $i$th coordinate equal to

$$\frac{\det(B_i)}{\det(A)}$$

**Chapter 3:** Vector Spaces

§1: **Basic concepts**

Basic idea: a vector space $V$ is a collection of things you can add together, and multiply by scalars (= numbers)

$V$ = things for which $v, w \in V$ implies $v + v \in V$ ; $a \in \mathbf{R}$ and $v \in V$ implies $a \cdot v \in V$

E.g., $V{=}\mathbf{R}^2$, add and scalar multiply componentwise

$V$=all 3-by-2 matrices, add and scalar multiply entrywise

$V{=}\{ax^2{+}bx{+}c : a, b, c \in \mathbf{R}\}$ = polynomials of degree $\leq$2; add, scalar multiply as functions

The *standard vector space* of dimension $n$ : $\mathbf{R}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbf{R} \text{ all } i\}$

An *abstract vector space* is a set $V$ together with some notion of addition and scalar multiplication, satisfying the 'usual rules': for $u, v, w \in V$ and $c, d \in \mathbf{R}$ we have

$u + v \in V$, $cu \in V$

$u + v = v + u$, $u + (v + w) = (u + v) + w$

There is $\mathbf{0} \in V$ and $-u \in V$ with $\mathbf{0} + u = u$ all $u$, and $u + (-u) = \mathbf{0}$

$c(u + v) = cu + cv$, $(c + d)u = cu + du$, $(cd)u = c(du)$, $1u = u$

Examples: $\mathbf{R}^{m,n}$ = all $m \times n$ matrices, under matrix addition/scalar mult

$C[a, b]$ = all continuous functions $f{:}[a, b]{\to}\,\mathbf{R}$, under function addition

$\{A \in \mathbf{R}^{n,n} : A^T = A\}$ = all symmetric matrices, is a vector space

Note: $\{f \in C[a, b] : f(a) = 1\}$ is **not** a vector space (e.g., has no bf 0)

Basic facts:

$0v = \mathbf{0}$, $c\mathbf{0} = \mathbf{0}$, $(-c)v = -(cv)$; $cv = \mathbf{0}$ implies $c = 0$ or $v = \mathbf{0}$

A vector space (=VS) has only one $\mathbf{0}$; a vector has only one additive inverse

Linear operators:

$T : V \to W$ is a linear operator if $T(cu + dv) = cT(u) + dT(v)$ for all $c, d \in \mathbf{R}$, $u, v \in V$

Example: $T_A : \mathbf{R}^n \to \mathbf{R}^m$, $T_A(v) = Av$, is linear

$T : C[a, b] \to \mathbf{R}$, $T(f) = f(b)$, is linear

$T : \mathbf{R}^2 \to \mathbf{R}$, $T(x, y) = x - xy + 3y$ is **not** linear!

§2: **Subspaces**

Basic idea: $V$ = vector space, $W \subseteq V$, then to check if $W$ is a vector space, using the **same** addition and scalar multiplication as $V$, we need only check **two things**:

whenever $c \in \mathbf{R}$ and $u, v \in W$, we **always** have $cu, u + v \in W$

All other properties come for free, since they are true for $V$ !

If $V$ is a VS, $W \subseteq V$ and $W$ is a VS using the same operations as $V$, we say that $W$ is a *(vector) subspace of* $V$.

Examples: $\{(x, y, z) \in \mathbf{R}^3 : z = 0\}$ is a subspace of $\mathbf{R}^3$

$\{(x, y, z) \in \mathbf{R}^3 : z = 1\}$ is **not** a subspace of $\mathbf{R}^3$

$\{A \in \mathbf{R}^{n,n} : A^T = A\}$ is a subspace of $\mathbf{R}^{n,n}$

Basic construction: $v_1, \cdots, v_n \in V$

$W = \{a_1v_1{+}\cdots a_nv_n : a_1, \ldots, a_n \in \mathbf{R}$ = all linear combinations of $v_1, \cdots, v_n$ = span$\{v_1, \cdots, v_n\}$

= the *span* of $v_1, \cdots, v_n$ , is a subspace of $V$

Basic fact: if $w_1, \ldots, w_k \in \text{span}\{v_1, \cdots, v_n\}$, then $\text{span}\{w_1, \cdots, w_k\} \subseteq \text{span}\{v_1, \cdots, v_n\}$

§3: **Subspaces from matrices**

column space of $A = \mathcal{C}(A) = \text{span}\{\text{the columns of } A\}$
row space of $A = \mathcal{R}(A) = \text{span}\{(\text{transposes of the }) \text{ rows of } A\}$
nullspace of $A = \mathcal{N}(A) = \{x \in \mathbf{R}^n : Ax = \mathbf{0}\}$

(Check: $\mathcal{N}(A)$ is a subspace!)

Alternative view $Ax$ = lin comb of columns of $A$, so is in $\mathcal{C}(A)$; in fact, $\mathcal{C}(A) = \{Ax : x \in \mathbf{R}^n\}$

Subspaces from linear operators: $T : V \to W$
image of $T = \text{im}(T) = \{Tv : v \in V\}$
kernel of $T = \ker(T) = \{x : T(x) = \mathbf{0}\}$

When $T = T_A$, $\text{im}(T) = \mathcal{C}(A)$, and $\ker(T) = \mathcal{N}(A)$

$T$ is called *one-to-one* if $Tu = Tv$ implies $u = v$
Basic fact: $T$ is one-to-one iff $\ker(T) = \{\mathbf{0}\}$

§4: **Norm and inner product**

Norm means length! In $\mathbf{R}^n$ this is computed as $||x|| = ||(x_1, \ldots, x_n)|| = (x_1^2 + \cdots + x_n^2)^{1/2}$

Basic facts: $||x|| \geq 0$, and $||x|| = 0$ iff $x = \mathbf{0}$,
$||cu|| = |c| \cdot ||u||$, and $||u + v|| \leq ||u|| + ||v||$ (triangle inequality)

unit vector: the norm of $u/||u||$ is 1; $u/||u||$ is the *unit vector* in the direction of $u$.
convergence: $u_n \to u$ if $||u_n - u|| \to 0$

Inner product:
idea: assign a number to a pair of vectors (think: angle between them?)
In $\mathbf{R}^n$, we use the *dot product*: $v = (v_1, \ldots, v_n)$, $w = (w_1, \ldots, w_n)$
$v \bullet w = \langle v, w \rangle = v_1 w_1 + \cdots + v_n w_n = v^T w$

Basic facts:
$\langle v, v \rangle = ||v||^2$ (so $\langle v, v \rangle \geq 0$, and equals 0 iff $v = \mathbf{0}$)
$\langle v, w \rangle = \langle w, v \rangle$; $\langle cv, w \rangle = \langle v, cw \rangle = c\langle v, w \rangle$

§5: **Applications of norms and inner products**

Cauchy-Schwartz inequality: for all $v, w$, $|\langle v, w \rangle| \leq ||v|| \cdot ||w||$
(this implies the triangle inequality)
So: $-1 \leq \langle v, w \rangle / (||v|| \cdot ||w||) \leq 1$

Define: the *angle* $\Theta$ between $v$ and $w$ = the angle (between 0 and $\pi$ with $\cos(\Theta) = \langle v, w \rangle / (||v|| \cdot ||w||)$
Ex: $v = w$ : then $\cos(\Theta) = 1$, so $\Theta = 0$
Two vectors are *orthogonal* if their angle is $\pi/2$, i.e., $\langle v, w \rangle$=0. Notation: $v \perp w$
Pythagorean theorem: if $v \perp w$, then $||v + w||^2 = ||v||^2 + ||w||^2$

Orthogonal projection: Given $v, w \in \mathbf{R}^n$, then we can write $v = cw + u$, with $u \perp w$

$c = \dfrac{\langle v, w \rangle}{\langle w, w \rangle}$;

$cw = \text{proj}_w v = \dfrac{\langle v, w \rangle}{\langle w, w \rangle} w = \dfrac{\langle v, w \rangle}{||w||} \dfrac{w}{||w||}$ = (orthogonal) projection of $v$ onto $w$

$u = v - cw$ !

Least squares:
Idea: Find the closest thing to a solution to $Ax = b$, when it ȟas no solution.
Overdetermined system: more equations than unknowns. Typically, the system will have no solution.
Instead, find the čl̆osest vector with a solution (i.e, of the form $Ax$) to b.
Need: $Ax - b$ perpendicular to the subspace $\mathcal{C}(A)$
I.e, need: $Ax - b \perp$ each column of $A$, i.e., need $\langle$(column of $A$),$Ax - b\rangle = 0$
I.e., need $A^T(Ax - b) = \mathbf{0}$, i.e., need $(A^TA)x = (A^Tb)$

Fact: such a system of equations is **always** consistent!
$Ax$ will be the closest vector in $\mathcal{C}(A)$ to $b$
**If** $A^TA$ is invertible (need: $r(A)$=number of columnsof $A$), then we can write $x = (A^TA)^{-1}(A^Tb)$; $Ax = A(A^TA)^{-1}(A^Tb)$

§6: **Bases and dimension**
Idea: putting free and bound variables on a more solid theoretical footing
We've seen: every solution to $Ax = b$ can be expressed in terms of the free variables $(x = v + x_{i_1}v_1 + \cdots + x_{i_k}v_k)$
Could a different method of solution give us a different number of free variables? (Ans: No! B/c that number is the 'dimension' of a certain subspace...)

Linear independence/dependence:
$v_1, \ldots, v_n \in V$ are linearly independent if the **only** way to express $\mathbf{0}$ as a linear combination of the $v_i$'s is with all coefficients equal to 0;
whenever $c_1v_1 + \cdots + c_nv_n = \mathbf{0}$, we have $c_1 = \cdots = c_n = 0$
**Otherwise**, we say the vectors are linearly dependent. I.e, some non-trivial linear combination equals $\mathbf{0}$. Any vector $v_i$ in such a linear combination having a non-zero coefficient is called **redundant**; the expression (lin comb $= \mathbf{0}$) can be rewritten to say that $v_i =$ lin comb of the remaining vectors, i.e., $v_i$ is in the **span** of the remaining vectors. This means: Any redundant vector can be removed from our list of vectors **without changing the span** of the vectors.

A **basis** for a vector space $V$ is a set of vectors $v_1, \ldots, v_n$ so that (a) they are linearly independent, and (b) $V$=span$\{v_1, \ldots, v_n\}$ .
Example: The vectors $e_1 = (1, 0, \ldots, 0)me_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1)$ are a basis for $\mathbf{R}^n$, the *standard basis.*

To find a basis: start with a collection of vectors that span, and repeatedly throw out redundant vectors (so you don't change the span) until the ones that are left are linearly independent. Note: each time you throw one out, you need to ask: are the remaining ones lin indep?
Basic fact: If $v_1, \ldots, v_n$ is a bassis for $V$, then every $v \in V$ can be expressed as a linear combination of the $v_i$'s in *exactly one way.* If $v = a_1v_1 + \cdots + a_nv_n$, we call the $a_i$ the **coordinates** of $v$ with respect to the basis $v_1, \ldots, v_n$ .

The Dimension Theorem: Any two bases of the same vector space contain the same number of vectors. (This common number is called the *dimension* of $V$, denoted dim$(V)$ .)
Reason: if $v_1, \ldots, v_n$ is a basis for $V$ and $w_1, \ldots, w_k \in V$ are linearly independent, then $k \leq n$

As part of that proof, we also learned:
If $v_1, \ldots, v_n$ is a basis for $V$ and $w_1, \ldots, w_k$ are linearly independent, then the spanning set $v_1, \ldots, v_n, w_1, \ldots, w_k$ for $V$ can be thinned down to a basis for $V$ by throwing away $v_i$'s .

**In reverse:** we can take any linearly independent set of vectors in $V$, and **add** to it from any basis for $V$, to produce a new basis for $V$.

Some consequences:
If $\dim(V)=n$, and $W \subseteq V$ is a subspace of $V$, then $\dim(W) \leq n$

If $\dim(V)=n$ and $v_1, \ldots, v_n \in V$ are linearly independent, then they also span $V$
If $\dim(V)=n$ and $v_1, \ldots, v_n \in V$ span $V$, then they are also linearly independent.

§7: **Linear systems revisited**
Using our new-found terminology, we have:

A system of equations $Ax = b$ has a solution iff $b \in \mathcal{C}(A)$ .

If $Ax_0 = b$, then every other solution to $Ax = b$ is $x = x_0 + z$, where $z \in \mathcal{N}(A)$ .

To finish our description of (a) the vectors $b$ that have solutions, and (b) the set of solutions to $Ax = b$, we need to find (useful) bases for $\mathcal{C}(A)$ and $\mathcal{N}(A)$.

So of course we start with:
Finding a basis for the row space.
Basic idea: if $B$ is obtained from $A$ by elementary row operations, then $\mathcal{R}(A) = \mathcal{R}(B)$.
So of $R$ is the reduced row echelon form of $A$, $\mathcal{R}(R) = \mathcal{R}(A)$
But a basis for $\mathcal{R}(R)$ is easy to find; take all of the non-zero rows of $R$ ! (The zero rows are clearly redundant.) These rows are linearly independent, since each has a 'special coordinate' where, among the rows, only it is non-zero. That coordinate is the *pivot* in that row. So in any linear combination of rows, only that vector can contribute something non-zero to that coordinate. *Consequently*, in any linear combination, that coordinate is the **coefficient** of our vector! **So**, if the lin comb is $\mathbf{0}$, the coefficient of our vector (i.e., each vector!) is 0.

Put bluntly, to find a basis for $\mathcal{R}(A)$, row reduce $A$, to $R$; the (transposes of) the non-zero rows of $R$ form a basis for $\mathcal{R}(A)$.

This in turn gives a way to find a basis for $\mathcal{C}(A)$, since $\mathcal{C}(A) = \mathcal{R}(A^T)$ !

To find a basis for $\mathcal{C}(A)$, take $A^T$, row reduce it to $S$; the (transposes of) the non-zero rows of $S$ form a basis for $\mathcal{R}(A^T) = \mathcal{C}(A)$ .

This is probably in fact the most useful basis for $\mathcal{C}(A)$, since each basis vector has that special coordinate. This makes it very easy to decide if, for any given vector $b$, $Ax = b$ has a solution. You need to decide if $b$ can be written as a linear combination of your basis vectors; but each coefficient will be the corrdinate of $b$ lying at the special coordinate of each vector. Then just check to see if **that** linear combination of your basis vectors adds up to $b$ !

There is another, perhaps less useful, but faster way to build a basis for $\mathcal{C}(A)$; row reduce $A$ to $R$, locate the pivots in $R$, and take the columns of $A$ (Note: $A$, **not** $R$ !) the correspond to the columns containing the pivots. These form a (different) basis for $\mathcal{C}(A)$.

Why? Imagine building a matrix $B$ out of just the bound columns. Then in row reduced form there is a pivot in every column. Solving $Bv = \mathbf{0}$ in the case that there are no free variables, we get $v = \mathbf{0}$, so the columns are linearly independent. If we now add a free column to $B$ to get $C$, we get the same collection of pivots, so our added column represents a free variable. Then there are non-trivial solutions to $Cv = \mathbf{0}$, so the columns of $C$ are not linearly independent. This means that the added columns can be expressed as a linear combination of the bound columns. This is true for all free columns, so the bound columns span $\mathcal{C}(A)$.

Finally, there is the nullspace $\mathcal{N}(A)$. To find a basis for $\mathcal{N}(A)$:

Row reduce $A$ to $R$, and use each row of $R$ to solve $Rx = \mathbf{0}$ by expressing each bound variable in terms of the frees. collect the coeeficients together and write $x = x_{i_1}v_1 + \cdots + x_{i_k}v_k$ where the $x_{i_j}$ are the free variables. Then the vectors $v_1, \ldots, v_k$ form a basis for $\mathcal{N}(A)$.

Why? By construction they span $\mathcal{N}(A)$; and just with our row space procedure, each has a special coordinate where only it is 0 (the coordinate corresponding to the free variable!).

Note: since the number of vectors in the bases for $\mathcal{R}(A)$ and $\mathcal{C}(A)$ is the same as the number of pivots ( = number of nonzero rows in the RREF) = rank of $A$, we have $\dim(\mathcal{R}(A))=\dim(\mathcal{C}(A))=r(A)$.

And since the number of vectors in the basis for $\mathcal{N}(A)$ is the same as the number of free variables for $A$ ( = the number of columns without a pivot) = nullity of $A$ (hence the name!), we have $\dim(\mathcal{N}(A)) = n(A) = n - r(A)$ (where $n$=number of columns of $A$).

So, $\dim(\mathcal{C}(A)) + \dim(\mathcal{N}(A))$ = the number of columns of $A$ .