Math 314 Topics for the second exam

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

Inverses.

Some conditions for/consequences of invertibility:

the following are all equivalent (A = n-by-n matrix).

0. A is invertible,
1. The RREF of A is I_n.
2. A has n pivots.
3. The equation $A\vec{x} = \vec{0}$ has only the solution x=0.
4. The columnsof A are linearly independent.
5. $T_A(\vec{x}) = A\vec{x}$ is one-to-one.
6. Every linear system $A\vec{x} = \vec{b}$ has a solution.
7. The columns of A span $\mathbb{R}^n$.
8. $T_A(\vec{x}) = A\vec{x}$ is onto.
9. There is a matrix C with $CA = I_n$.
10. There is a matrix B with $AB = I_n$.
11. A^T is invertible,
12. For one choice of $\vec{b}$, $A\vec{x} = \vec{b}$ has a unique solution.
13. For every choice of $\vec{b}$, $A\vec{x} = \vec{b}$ has a unique solution.

The equivalence of 3. and 13. is sometimes stated as **Fredholm's alternative**: Either every equation $A\vec{x} = \vec{b}$ has a unique solution, or the equation $A\vec{x} = \vec{0}$ has a **non-trivial** solution (and only one of the alternatives can occur).

Determinants.

(Square) matrices come in two flavors: invertible (all $Ax = b$ have a solution) and non-invertible ($Ax = \vec{0}$ 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 ith row and jth 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 $\det(A) \neq 0$.

In fact, we can **use** row operations to calculate $\det(A)$ (since the RREF of a matrix is upper triangular). We just need to *keep track* of the row operations we perform, and compensate for the changes in the determinant;
$\det(A) = (1/c)\det(E_i(c)A)$, $\det(A) = (-1)\det(E_{ij}A)$

More interesting facts:
$\det(AB) = \det(A)\det(B)$; $\det(A^T) = \det(A)$; $\det(A^{-1}) = [\det(A)]^{-1}$

We can expand along other columns than the first: for any fixed value of j (= column),
$\det(A) = \Sigma_{i=1}^{n}(-1)^{i+j}a_{ij}\det(M_{ij}(A))$
(expanding along jth column)
And since $\det(A^T) = \det(A)$, we could expand along **rows**, as well.... for any fixed i (= row),
$\det(A) = \Sigma_{j=1}^{n}(-1)^{i+j}a_{ij}\det(M_{ij}(A))$

Vector Spaces.

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 $v + v \in V$; $a \in \mathbb{R}$ and $v \in V$ implies $a \cdot v \in V$

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

V=all 3-by-2 matrices, add and scalar multiply entrywise
$\mathcal{P}_2$=$\{ax^2+bx+c : a,b,c \in \mathbb{R}\}$ = polynomials of degree ≤2; add, scalar multiply as functions
More generally: $\mathcal{P}_n = \{$all polynomials of degree $\le n\}$ is a vector space

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

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 $c,d \in \mathbb{R}$ we have

$u+v \in V$, $cu \in V$
$u+v=v+u$, $u+(v+w)=(u+v)+w$
There is $\vec{0} \in V$ and $-u \in V$ with $\vec{0}+u=u$ all u, and $u+(-u)=\vec{0}$
$c(u+v)=cu+cv$, $(c+d)u=cu+du$, $(cd)u=c(du)$, $1u=u$

Examples: $\mathbb{R}^{m,n}$ = all $m \times n$ matrices, under matrix addition/scalar mult
$C[a,b]$ = all continuous functions f:$[a,b] \to \mathbb{R}$, under function addition
$\{A \in \mathbb{R}^{n,n} : A^T = A\}$ = 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 $\vec{0}$)
Basic facts:
$0v=\vec{0}$, $c\vec{0}=\vec{0}$, $(-c)v=-(cv)$; $cv=\vec{0}$ implies $c=0$ or $v=\vec{0}$
A vector space (=VS) has only one $\vec{0}$; a vector has only one additive inverse

Linear operators/transformations:
$T: V \to W$ is a linear operator if $T(cu+dv)=cT(u)+dT(v)$ for all $c,d \in \mathbb{R}$, $u,v \in V$
Example: $T_A : \mathbb{R}^n \to \mathbb{R}^m$, $T_A(v)=Av$, is linear
$T: C[a,b] \to \mathbb{R}$, $T(f)=f(b)$, is linear

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

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 $c \in \mathbb{R}$ and $u,v \in W$, we **always** have $cu, u+v \in W$ (W is "closed" under addition and scalar multiplication).

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: $\{(x,y,z) \in \mathbb{R}^3 : z=0\}$ is a subspace of $\mathbb{R}^3$
$\{(x,y,z) \in \mathbb{R}^3 : z=1\}$ is **not** a subspace of $\mathbb{R}^3$
$\{A \in \mathbb{R}^{n,n} : A^T = A\}$ is a subspace of $\mathbb{R}^{n,n}$

Basic construction: $v_1,\cdots,v_n \in V$
$W = \{a_1v_1 + \cdots a_nv_n : a_1,\ldots,a_n \in \mathbb{R}$ = 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 $\text{span}\{w_1,\cdots,w_k\} \subseteq \text{span}\{v_1,\cdots,v_n\}$

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

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

Alternative view Ax = lin comb of columns of A, so is in col(A); in fact, col(A) = $\{Ax : x \in \mathbb{R}^n\}$. So col($A$) is the set of vectors b for which $Ax=b$ has a solution. Any two solutions $Ax=b=Ay$ have $A(x-y)=AX-Ay=b-b=0$, so $x-y$ is in null(A). So the collection of <u>all</u> solutions to $AX=b$ are (particular solution)+(vector in null(A)). So col(A) tells is which SLEs have solutions, and null(A) tells us how many solutions there are.

The descriptions of $\mathcal{C}(A)$ and $\mathcal{N}(A)$ are fundamentally different; $\mathcal{C}(A)$ tells us how to (quickly) <u>build</u> elements of the subspace, while $\mathcal{N}(A)$ tells us how to (quickly) <u>decide</u> if a vector is in the subspace. As written, doing the other thing for the other subspace requires 'work' (that is, row reduction!). But we can (by row reduction) describe each subspace in terms matching the other, to make the corresponding task - building versus deciding - quick, as well.

To view $\mathcal{N}(A)$ as a column space, row reduce A! In RREF, we can write the solutions to $A\vec{x}=\vec{0}$ as a linear combination of vectors, one for each free variable, by solving for each pivot variable in terms of the free ones. These vectors then span $\mathcal{N}(A)$; writing them as the columns of a matrix B, we have $\mathcal{N}(A)=\mathcal{C}(B)$.

To view $\mathcal{C}(A)$ as a nullspace, row reduce the super-augmented matrix $(A|I_n)\to(R|B)$. From our inverse work, we know that $BA=R$. So $(A|\vec{b})$ row reduces to $(R|B\vec{b})$. In RREF, the rows r_i of B opposite the rows of 0's in R tell us that for $(A|\vec{b})$ to be consistent we must have $(r_i)\vec{b}=0$ for each i. Assembling these r_i into a matrix, Q, we then have that $\vec{b}\in\mathcal{C}(A)$ precisely when $Q\vec{b}=\vec{0}$, i.e., $\vec{b}\in\mathcal{N}(Q)$. So $\mathcal{C}(A)=\mathcal{N}(Q)$!

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

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 if and only if $\ker(T)=\{\vec{0}\}$

Bases, dimension, and rank

A **basis** for a subspace W of V is a set of vectors $v_1,\ldots,v_n\in W$ so that (a) they are linearly independent, and (b) W=span$\{v_1,\ldots,v_n\}$.
Example: The vectors $e_1=(1,0,\ldots,0), e_2=(0,1,0,\ldots,0),\ldots,e_n=(0,\ldots,0,1)$ are a basis for $\mathbb{R}^n$, the *standard basis*.

The idea: a basis allows you to express every vector in the subspace as a linear combination <u>in exactly one way</u>.

A system of equations $Ax=b$ has a solution iff $b\in$col(A) .

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

The row, column, and nullspaces of a matrix A are therefore useful spaces (they tell us useful things about solutions to the corresponding linear system), so it is useful to have bases for them.

Finding a basis for the row space.
Basic idea: if B is obtained from A by elementary row operations, then row(A) =row(B).
So of R is the reduced row echelon form of A, row(R) =row(A)
But a basis for row(R) is quick to identify; 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 $\vec{0}$, the coefficient of our vector (i.e., of each vector!) is 0.

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

This in turn gives a way to find a basis for col(A), since col(A) =row(A^T) !

To find a basis for col(A), take A^T, row reduce it to S; the (transposes of) the non-zero rows of S form a basis for row(A^T) =col(A) .

This is probably in fact the most useful basis for col(A), since each basis vector has that special coordinate. This makes it very quick 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 coordinate 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 col(A); row reduce A to R, locate the pivots in R, and take the columns of A (Note: A, **not** R !) that correspond to the columns containing the pivots. These form a (different) basis for col(A).

Why? Imagine building a matrix B out of just the pivot columns. Then in row reduced form there is a pivot in every column. Solving $Bv = \vec{0}$ in the case that there are no free variables, we get $v = \vec{0}$, 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 $Cv = \vec{0}$, 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 col(A).

Finally, there is the nullspace null(A). To find a basis for null(A):

Row reduce A to R, and use each row of R to solve $Rx = \vec{0}$ by expressing each bound variable in terms of the frees. collect the coefficients 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 null(A).

Why? By construction they span null(A); and just as with our row space procedure, each has a special coordinate where only it is not 0 (the coordinate corresponding to the free variable!).

More on Bases.

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 basis 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 $v_1, \ldots, v_n$. We can then <u>think</u> of v as the vector $(a_1, \ldots a_n)^T$ = the <u>coordinates</u> of v with respect to the basis $v_1, \ldots, v_n$, so we can think of V as "really" being $\mathbb{R}^n$.

The Basis 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.

Rank of a matrix = r(A) = number of non-zero rows in RREF = number of pivots in RREF = dim(col(A)).
Nullity of a matrix = n(A) = number of columns without a pivot = # columns − # pivots = dim(null(A))
rank = number of bound variables, nullity = number of free variables
rank $\leq$ number of rows, number of columns (at most one pivot per row/column!)

Note: since the number of vectors in the bases for row(A) and col(A) is the same as the number of pivots (= number of nonzero rows in the RREF) = rank of A, we have dim(row(A))=dim(col(A))=$r(A)$.
And since the number of vectors in the basis for null(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(null(A)) = $n(A)$ = $n - r(A)$ (where n=number of columns of A).
So, dim(col(A)) + dim(null(A)) = the number of columns of A .
rank + nullity = number of columns = number of variables

Using coordinates, we can treat any vector space as $\mathbb{R}^k$ for some k, and so carry out these kinds of computations - find a basis inside a spanning set, test for linear independence, test for spanning, extend a linearly independent set to a basis, etc. For example, since the polynomials $\{1, x, x^2, x^3, x^4\}$ are a basis for $\mathcal{P}_4$, we can test the linear independence of $\{x + 2, x^2 - x - 3, 2x^2 + x\}$ (they aren't) by testing their coordinate vectors.