MATH 310, REVIEW SHEET 2

These notes are a very short summary of the key topics in the book (and follow the book pretty closely). You should be familiar with everything on here, but it's not comprehensive, so please be sure to look at the book and the lecture notes as well. Let me know if you notice any particularly egregious omissions and I'll add more.

These notes also don't contain any examples at all. If you run into something unfamiliar, look at the notes or the book for an example! The goal of this sheet is just to remind you of what the major topics are.

I've also indicated some of the important "problem types" we've encountered so far and that you should definitely be able to do. There will inevitably be a problem on the exam not of one of the types listed.

2. Matrix algebra

2.5. **Matrix factorizations and LU decomposition.** The rest of chapter 2 was on the list midterm, but there is one section to brush up on.

If A is an $m \times n$ matrix, one can (usually) find an LU decomposition $A = LU$. Here L is an $m \times m$ lower triangular matrix (note: the size is different from that of A, and L is always square). The way we arrange things, L is always going to have 1s on the diagonal. The matrix U is upper triangular, and the same size as A.

Here's how to find it: start with your matrix A, and do row reduction until you reach echelon form (NB: not rref. However, you're only allowed to use certain row operations: you can subtract a multiply of a row from a row below it (or equivalently, add a multiple of a row to a row below it). But you can't multiply a row by a number, and you can't swap rows. (Well, really you can, but you need to use a slightly more complicated form of LU decomposition if you want to, and we didn't cover it.)

The matrix U is just going to be the echelon form that you reached. To get L, start off with an $m \times m$ square matrix. Put 1s along the main diagonal and 0s above it. Now, in position (i, j), put the number of Row j's that you subtracted from Row i in the course of doing row reduction. For example, L_{32} is the number of row 2's you subtracted from row 3.

LU decomposition essentially "remembers" how to do row reduction on a matrix A. This has a number of computational advantages. If you want to solve $A\mathbf{x} = \mathbf{b}$, you can instead solve $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$, and then $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$. It may not sound like it, but if you are going to need to solve $A\mathbf{x} = \mathbf{b}$ for many different $\mathbf{b}$'s for a fixed matrix A, this is going to be faster. The point is that $L\mathbf{x} = \mathbf{y}$ only involves doing row reduction on a triangular matrix, which is fast.

- Find LU decomposition of a given matrix.
- Use this to solve $A\mathbf{x} = \mathbf{b}$ by solving $L\mathbf{y} = \mathbf{b}$ and then $U\mathbf{x} = \mathbf{y}$.

Revision e75a09c, jdl, 2015-11-09

3. Determinants

3.1. Introduction to determinants.

3.2. Properties of determinants.

If A is an $n \times n$ matrix (NB: it has to be square), then $\det A$ (also written $|A|$) is a number. Be sure you know the following useful properties of this operation:

(1) Determinant is unchanged when adding one row/column to another row/column.
(2) Determinant changes sign when swapping two rows/columns.
(3) $\det A = 0$ if and only if A has a nonzero nullspace (i.e. it's not invertible).

We have three main methods for computing determinants. I'm not going to explain how each goes, but make sure you know them all. Here they are, with some suggestions one when each one is useful.

(1) Cofactor expansion. Especially suitable when the matrix has a row or column with only a couple nonzero entries. You can always start off with cofactor expansion and then switch to another method for the smaller determinants that pop out.
(2) Product of pivots. If a matrix doesn't fit any of the above (i.e. it's at least 4×4 and doesn't have many zeroes), this is probably the way to go. Do row reduction, without ever multiplying a row by a number. When you get to echelon form, the determinant is just the product of the pivots, with an extra factor of -1 for every time you had to swap rows.
(3) The specific formulas for 2×2 and 3×3 matrices. Good if you are dealing with a 2×2 and 3×3 matrix; obviously not so helpful otherwise.

Perhaps the following should qualify as a basic methods:

(1) Triangular matrices: just multiply the diagonal entries.
(2) Reduce to simpler matrix by row/column operations. These don't change the determinant, so if you can do something sneaky like reduce to an triangular matrix or a matrix with lots of 0's using just a couple row/column operations (including row swaps), you're in business.

- Find the determinant of a given matrix by whatever method is appropriate.

3.3. Cramer's rule, volume, and linear transformations.

The determinant of A is equal to the volume of the parallelpiped defined by the columns of A; this shows up as the Jacobian in change of coordinates for 2- or 3-dimensional integrals. A consequence of this is that if you look at the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by the matrix A, it rescales area/volume by a factor of the determinant.

Cramer's rule is a way to solve linear systems $A\mathbf{x} = \mathbf{b}$, where A is square. Here's what you do. Let $A_i(\mathbf{b})$ be the matrix obtained by replacing the i^{th} column of A with the vector $\mathbf{b}$. The $x_i = \det A_i(\mathbf{b}) / \det A$.

As a consequence, we get a formula for A^{-1}: we have $A^{-1} = C^T / \det A$, where C is the matrix of cofactors of A. There are examples of this in the lecture notes, book, and homework solutions, but there hasn't been one on a quiz. You might want to take a look.

These are not efficient ways to solve $A\mathbf{x} = \mathbf{b}$ in general, since you have to take a lot of determinants, which isn't fun. Nevertheless it can sometimes be very helpful to have an explicit formula like this; for example, when your matrix A involves a parameter s, and you want to understand how the solution depends on the parameter s, Cramer's rule gives you a good way to go about it.

- Use Cramer's rule to solve $A\mathbf{x} = \mathbf{b}$, especially when A includes parameters.
- Find the area of a parallelograph with given vertices, using determinants.
- Find the entries of the inverse of a 3×3 matrix using Cramer's rule.

4. Vector spaces

4.1. **Vector spaces and subspaces.** A vector space is any bunch of things that you can add and multiply by scalars in a way that satisfies the usual rules of arithmetic. This includes both normal old vectors $\mathbb{R}^n$ as well as other examples. The one we talked about the most was $\mathbb{P}^n$, the polynomials of degree at most n.

If you have a vector space, a subspace of that vector space is a collection of vectors such that if you add any two, you get another, and if you multiply one by a scalar, you get another. We talked about quite a few examples and non-examples; check the notes to see some.

One basic example of a subspace is the following: if you have a bunch of vectors $\mathbf{v}_1, \dots, \mathbf{v}_r$ in $\mathbb{R}^n$, then their span (i.e. the set of all linear combinations of the vectors) is a subspace.

- Tell if a given collection of vectors is a subspace or not.

4.2. **Null spaces, column spaces, and linear transformations.**

4.3. **Linearly independent sets; bases.** Suppose you have an $m \times n$ matrix A. It determines two very important subspaces: the null space $\operatorname{Nul} A$ and the column space $\operatorname{Col} A$.

The column space of A is just the span of the columns of A. All the columns are vectors in $\mathbb{R}^m$, so the column space is a subspace of $\mathbb{R}^m$. The null space of A is the set of all solutions to $A\mathbf{x} = \mathbf{0}$; this is just the set of solutions to a homogeneous linear system. This is a subspace of $\mathbb{R}^n$.

A bunch of vectors inside a subspace (or a vector space) is called a *basis* if the vectors are linearly independent and they span the subspace. This means that every vector in the subspace is a combination of vectors in the basis. But the basis can't have too many vectors in it: then they wouldn't be linearly independent.

You should know how to find a basis for the nullspace of a matrix, and a basis for the column space of a matrix. This isn't so hard: for the nullspace, what you need to do is find the general solution of $A\mathbf{x} = \mathbf{0}$, in parametric vector form. Your basis for the nullspace is then given by the vectors that appear in parametric vector form; you'll end up with one for each free variable. To find a basis for the column space of A, run row reduction on A until you reach an echelon matrix, U. Figure out which columns of U are the pivot columns. Your basis for the column space is then given by the corresponding columns of the original matrix A (not of U!)

Don't lose sight of what all this means. A basis for the column space is a collection of vectors such that everything in the nullspace (i.e. every solution of $A\mathbf{x} = \mathbf{0}$ is a linear combination of those vectors in a unique way). A basis for the column space is a collection of linearly independent vectors that span the column space – essentially what our procedure does is get rid of all the columns of A that are linearly dependent on the preceding columns, leaving us with a linearly independent set.

- Know what a basis is, and be able to check whether a given set of vectors is a basis for $\mathbb{R}^n$.
- Find a basis for the null space and column space of a matrix A.
- Find a basis for the span of a given set of vectors.

4.4. **Coordinate systems.** The important thing about a basis for a subspace is that every vector is a combination of the vectors in a unique way. What this means is that every x in the subspace can be written as

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$$

in only one way. The numbers c_i are called the coordinates of x with respect to the basis $\mathcal{B}$, and we write $[\mathbf{x}]_\mathcal{B}$ for the vector whose entries are the c_i's.

If you know $[\mathbf{x}]_\mathcal{B}$ for some basis $\mathcal{B}$, you can easily figure out what $\mathbf{x}$ is: it's just $\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3$. In matrix form, we can write this as $\mathbf{x} = \mathcal{P}_\mathcal{B}([\mathbf{x}]_\mathcal{B})$. Here $\mathcal{P}_\mathcal{B}$ is the "coordinate matrix": it's just the matrix you get when you write the vectors in your basis as columns.

The other thing to do is to be able to find $[\mathbf{x}]_\mathcal{B}$ if you're given $\mathbf{x}$. This means you want to write $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 = \mathbf{x}$, which is just a vector equation in c_1 and c_2. You can find them by writing down the augmented matrix and doing row reduction. An alternative is to use the formula above to obtain $\mathbf{x} = \mathcal{P}_\mathcal{B}^{-1}([\mathbf{x}]_\mathcal{B})$ (the downside to this is that you have to invert a matrix to use it).

- Find $\mathbf{x}$ given $[\mathbf{x}]_\mathcal{B}$.
- Find $[\mathbf{x}]_\mathcal{B}$ given $\mathbf{x}$.
- Work with coordinates for vectors in $\mathbb{P}_n$, the set of polynomials of degree $\leq n$.

4.5. **The dimension of a vector space.** The dimension of a vector space (or subspace of a vector space) is the number of vectors in a basis. The reason this makes sense is that any two bases have the same number of vectors – otherwise this wouldn't be a very useful definition.

To find the dimension of a vector space, you just have to find a basis (using one of the methods discussed above), and then count how many vectors you ended up with. That's all there is to it.

- Find the dimension of $\operatorname{Nul} A$ for a matrix A.
- Find the dimension of $\operatorname{Col} A$ for a matrix A.

4.6. **Rank.** The rank of a matrix is the dimension of its column space. The main fact about it is *the rank theorem*, which says that if A is an $m \times n$ matrix,

$$\operatorname{rank} A + \dim \operatorname{Nul} A = n.$$

This lets you find the dimension of the null space if you know the rank, and vice versa. This is surprisingly handy; there were a few questions on the homework about this.

Another topic introduced in this section was the row space of a matrix. As you can probably guess by now, the row space of an $m \times n$ matrix is the span of its rows. Since the rows have n entries, this is a subspace of $\mathbb{R}^n$. There's a recipe to compute it, as usual: do row reduction on A until you reach an echelon form U. A basis for the row space is given by the nonzero vectors in U (remember that your echelon form is likely to have a bunch of rows of 0s at the bottom; we don't want those). The other important fact about rank is that the dimension of the row space is equal to the dimension of the column space. (This is not obvious at all!)

A note: given a matrix A, you can find bases for the row space and column space just by doing row reduction until you get to echelon form. Then you see what are the pivot columns (which tells you the column space) and what are the nonzero rows (which tells you the row space).

- Compute the rank of a given matrix.
- Find the rank of a matrix given the dimension of the null space, or the other way a round.
- Find a basis for the row space of a matrix.

4.7. **Change of basis.** This is possibly the most annoying chapter, but it is important to know. Sometimes you are dealing with some vector space, and you have two different bases $\mathcal{B}$ and $\mathcal{C}$. It's important to know how to find the coordinates of a vector in $\mathcal{C}$ fi you already know the coordinates in $\mathcal{B}$, and vice versa. It turns out that all you need to do is apply a certain matrix to the coordinates. More precisely,

$$[\mathbf{x}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathcal{P}}([\mathbf{x}]_{\mathcal{B}}).$$

The question is how to figure out what the matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathcal{P}}$ is. There are two versions of this, depending on what exactly we're dealing with.

Possibly the main case of this is when you have two bases $\mathcal{B}$ and $\mathcal{C}$ for $\mathbb{R}^n$. In this case, we the formula is pretty easy: you want $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathcal{P}} = \mathcal{P}_{\mathcal{C}}^{-1}\mathcal{P}_{\mathcal{B}}$. You can either calculate this directly (pretty painless in the 2×2 case), or use a row reduction trick. Start with $[\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{b}_1\ \mathbf{b}_2]$, do row reduction, and you end up with $[I\ A]$, and A is the matrix you want to use $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathcal{P}} = [[\mathbf{b}_1]_{\mathcal{C}}\ [\mathbf{b}_2]_{\mathcal{C}}\ [\mathbf{b}_3]_{\mathcal{C}}]$: express each of your $\mathcal{B}$-basis vectors in $\mathcal{C}$-coordinates, and use that as the change of basis matrix.

The other version of this is when you have two bases for a vector space that isn't $\mathbb{R}^n$; maybe it's the column space of a matrix, or maybe it's a set of polynomials,. . . . In that case you want

Another thing to know: if you want to go from $\mathcal{C}$ to $\mathcal{B}$ instead of the other way around, you can use the fact that $\underset{\mathcal{B}\leftarrow\mathcal{C}}{\mathcal{P}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathcal{P}}^{-1}$.

- Find the change of basis matrix (in either situation above).
- Use it to convert coordinates of vectors between different coordinate systems.

4.9. **Application to Markov chains.** A matrix A is called a *stochastic* matrix if all its entries are non-negative and the columns all add up to 1. A vector is called a probability vector if its entries add up to 1. If you apply a stochastic matrix to a probability vector, you get another probability vector. A *Markov chain* is a sequence of probability vectors obtained by taking $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

We ran into some examples earlier in the course, in Section 1.9 when we discussed migration matrices. There are lots more examples in the book and the homework. make you you understand them, and you can translate a word problem into the formula for a Markov process.

The part that's new here was about "steady-state vectors". The deal is this: if you apply a stochastic matrix A to a vector many, many, many times, the vectors $\mathbf{x}_k$ you get are going to get very close to a certain vector $\mathbf{q}$, called the *steady state*. This is a kind of equilibrium: if the system is describing migration, it will eventually level off. The amounts that it levels off at are the steady states.

Finding the steady state isn't so hard. We want a vector $\mathbf{q}$ with $A\mathbf{q} = \mathbf{q}$. That means that $(A - I)\mathbf{q} = \mathbf{0}$. So write down the matrix $A - I$, and find $\mathbf{q}$ that's in its nullspace. That will be the steady state you're after.

- Translate a word problem into a stochastic matrix.
- Find the steady state of a Markov chain.

5. Eigenvalues and eigenvectors

5.1. Eigenvectors and eigenvalues.

5.2. The characteristic equation.

A vector $\mathbf{x}$ is called an eigenvector of the matrix A if $A\mathbf{x} = \lambda\mathbf{x}$, where λ is a number (called the eigenvalue): this means that $A\mathbf{x}$ points in the same direction as $\mathbf{x}$ (or the opposite direction, in case $\lambda < 0$).

This is equivalent to saying that $(A - \lambda I)\mathbf{x} = \vec{0}$, which is to say that $\mathbf{x}$ is in the nullspace of $A - \lambda I$. This is the observation that lets us find the eigenvectors.

There are a couple cases where this has particular geometric significance: $\mathbf{x}$ is an eigenvector with $\lambda = 1$ means that $\mathbf{x}$ doesn't change when you apply A; with $\lambda = 0$ means that $A\mathbf{x} = \vec{0}$, i.e. $\mathbf{x}$ is in the nullspace of A; with $\lambda = -1$ means that the direction of $\mathbf{x}$ is reversed when we apply A to $\mathbf{x}$.

For most values of λ, the matrix $A - \lambda I$ won't have a nullspace at all. The only times it does is when $\det(A - \lambda I) = 0$, and so the eigenvalues are precisely the solutions of $\det(A - \lambda I) = 0$. Once you've find an eigenvalue, the way to find the corresponding eigenvector is to write down the matrix $A - \lambda I$ for that value of λ, and then find something in the nullspace using the usual procedure (elimination, special solutions).

- Find the eigenvalues of a matrix.
- Find the eigenvectors corresponding to all the eigenvalues.

5.3. Diagonalization.

If we have a bunch of eigenvectors for a matrix A, we can put all of them as the columns of a matrix P, and the eigenvalues as the diagonal entries of a diagonal matrix D. These will satisfy $A = PDP^{-1}$. Part of doing diagonalization is knowing how to invert the matrix P, something we covered earlier in the course, so be ready for it. Make sure you remember the quick way to invert a 2×2 matrix. This will let you find P^{-1} without having to think too hard.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

One thing you can do with this is compute powers of a matrix with ease. If we have $A = PDP^{-1}$, then $A^n = (PDP^{-1})^n = PD^nP^{-1}$. To find D^n you just raise the diagonal entries of D to the nth power. Then to find A^n, just multiply out PD^nP^{-1}, a product of three matrices.

Not every matrix is diagonalizable. If A has n *distinct* eigenvalues, then for each eigenvalue we can find an eigenvector. Eigenvectors with different eigenvalues are automatically independent, so that gives us n independent eigenvectors. Put those into a matrix P as above, and tada, it's diagonalized. Let me stress: if the eigenvalues are all distinct, diagonalization is automatic. If there's a repeated eigenvalue, things can go either way: maybe it's diagonalizable, maybe it isn't. You have to check.

Problems can arise when A has a repeated eigenvalue. It's only guaranteed that we can find a single eigenvector for that eigenvalue, which isn't enough to make a square matrix P.

It's still possible that A can be diagonalized, but you actually need to check for eigenvectors. The typical example of this is something like $\left(\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right)$, which has only a single eigenvector.

- Compute the diagonalization of a matrix A.
- Compute powers A^n of a matrix A.