DATA605

Week3 and Week4 Lecture

Vector Spaces

A real vector space is a set of vectors together with rules for vector addition and multiplication by real numbers. Addition and multiplication must produce vectors in the space, and follow these properties:

1. $x + y = y + x$.
2. $x + (y + z) = (x + y) + z$.
3. There is a unique "zero vector" such that $x + 0 = x$ for all x.
4. For each x there is a unique vector $-x$ such that $x + (-x) = 0$.
5. $1x = x$.
6. $(c_1 c_2)x = c_1(c_2 x)$.
7. $c(x + y) = cx + cy$.
8. $(c_1 + c_2)x = c_1 x + c_2 x$.

Vector Spaces and Subspaces

A subspace of a vector space is a nonempty subset that satisfies the requirements for a vector space:

- Linear combinations stay in the subspace.
- Geometrically, think of the usual three-dimensional R 3 and choose any plane through the origin. That plane is a vector space in its own right.
- Notice in particular that the zero vector will belong to every subspace

Column Space

We now come to the key examples, the column space and the nullspace of a matrix **A**. The **column space** contains all linear combinations of the columns of **A**. It is a subspace of $\mathbf{R}^m$.

- In some sense, the column space is the ‘range’ of vectors than can come from matrix **A** being multiplied by vector **x**. Or possible values of **b** in the **Ax = b** equation

Null Space

The second approach to **Ax = b** is “dual” to the first. We are concerned not only with attainable right-hand sides b, but also with the solutions x that attain them. The right-hand side **b = 0** always allows the solution **x = 0**, but there may be infinitely many other solutions. (There always are, if there are more unknowns than equations, $n > m$.)

The solutions to **Ax = 0** form a vector space—the nullspace of **A**. The nullspace of a matrix consists of all vectors x such that **Ax = 0**. It is denoted by **N(A)**. It is a subspace of $\mathbf{R}^n$, just as the column space was a subspace of $\mathbf{R}^m$.

We want to be able, for any system Ax = b, to find C (A) and N (A): all attainable right-hand sides b and all solutions to Ax = 0.

Nullspace and Rank

(pg93) For an invertible matrix, the nullspace contains only $x = 0$ (multiply $Ax = 0$ by A^{-1}). The column space is the whole space ($Ax = b$ has a solution for every b). The new questions appear when the nullspace contains more than the zero vector and/or the column space contains less than all vectors:

Elimination reveals the pivot variables and free variables. If there are **r pivots**, there are **r pivot variables** and **n − r free variables**. That important number r will be given a name—it is the **rank** of the matrix.

Linear Independence

Suppose $c_1 v_1 + \cdots + c_k v_k = 0$ only happens when $c_1 = \cdots = c_k = 0$. Then the vectors $v_1, \ldots, v_k$ are linearly independent. If any c’s are nonzero, the v’s are linearly dependent. One vector is a combination of the others.

A simpler way of saying this is that a set of vectors are linearly independent as long as each vector cannot be seen as a linear combination of the others.

A set of n vectors in R^m must be linearly dependent if $n > m$.

Why do we care? How does it connect to spaces?

The columns of A are independent exactly when N (A) = {zero vector}. To check any set of vectors $v_1, \ldots, v_n$ for independence, put them in the columns of A. Then solve the system $Ac = 0$; the vectors are dependent if there is a solution other than $c = 0$. With no free variables (rank n), there is no nullspace except $c = 0$;

Spanning a Space

What does it mean to span a space or subspace? Basically it means having a set of vectors that can describe all vectors in the space as a linear combination.

Formally: If a vector space V consists of all linear combinations of w 1 , . . . , w , then these vectors span the space. Every vector v in V is some combination of the w’s.

Informally: Think of span as like a coordinate system: The coordinate vectors e 1 , . . . , e n coming from the identity matrix span R n

Basis of Subspace

Combination of spanning and linear independence: A basis for **V** is a sequence of vectors having two properties at once:

1. The vectors are linearly independent (not too many vectors).

2. They span the space V (not too few vectors).

There is one and only one way to write v as a combination of the basis vectors. A vector space has infinitely many different bases.

Connection: The columns that contain pivots are a basis for the column space; the number of free variables gives the dimension of the nullspace of A.

Useful link: http://palmer.wellesley.edu/~aschultz/f07/math225/coursenotes_and_handouts/coursenotes_071023.pdf

Dimension

Any two bases for a vector space V contain the same number of vectors. This number, which is shared by all bases and expresses the number of “degrees of freedom” of the space, is the **dimension of V.**

The point is that a basis is a maximal independent set. It cannot be made larger without losing independence. A basis is also a minimal spanning set. It cannot be made smaller and still span the space.

Extra: Four Fundamental Spaces

The author points out that there are four fundamental spaces associated with a matrix A. We went over two of them. We also have spaces associated with the transpose of A.

The following picture shows how they connect.

For more information from the author of the book, checkout:

- https://www.youtube.com/watch?v=nHlE7EgJFds
- http://www.engineering.iastate.edu/~julied/classes/CE570/Notes/strangpaper.pdf
- http://web.mit.edu/18.06/www/Essays/newpaper_ver3.pdf

Four Fundamental Spaces

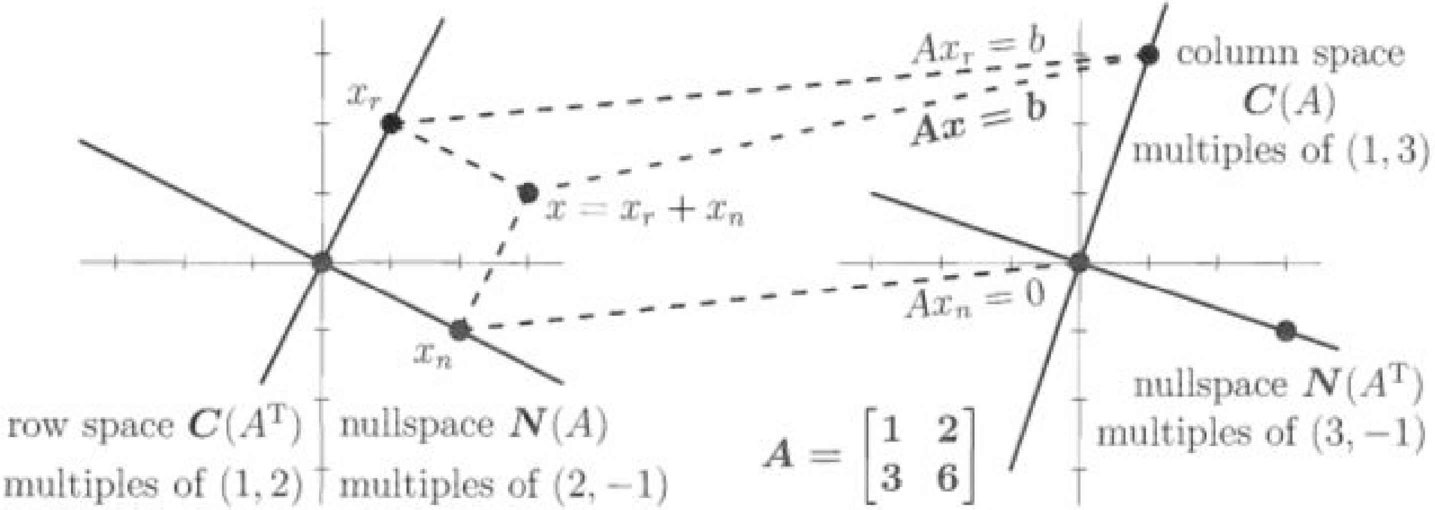

Figure 2.5: The four fundamental subspaces (lines) for the singular matrix A.

Determinants

The determinant of a matrix is one of the fundamental properties of a matrix. It is usually designated by long vertical lines surrounding a matrix: $|A|$. Here are some important things to know about this property:

1. The determinant is a real number.
2. The determinant can be a negative number.
3. The determinant only exists for **square** matrices.
4. The inverse of a matrix will exist only if the determinant is not zero.

Some other useful ones are:

1. Switching two rows or columns changes the sign (this is why we have alternating signs in one of the ways to calc)
2. Multiplying one row by a constant multiplies the whole determinant by that constant (think of how this might tie to volume)
3. If A is triangular then det A = product of diagonal entries.

Determinants, what are they really?

Ok, this is great but what exactly is this? Its tough to give a concrete, intuitive way to think about the determinant. However, the accepted answer at http://math.stackexchange.com/questions/668/whats-an-intuitive-way-to-think-about-the-determinant is a great resource. The TL;DR version:

The determinant of the linear transformation (matrix) T is the signed volume of the region gotten by applying T to the unit cube.

The main thing you need to know is how to compute it

Determinants: How to Compute

The week3 lecture notes go over the computation. The book, on page 259, goes over a few other ways, including finding and multiplying pivots and ‘the big formula’. To the right you see R code for finding a determinant

```
> A <- matrix(c(1,2,3,8,6,3,2,2,6), nrow=3, byrow=T)
> det(A)
[1] -42
> solve(A)
           [,1]        [,2]       [,3]
[1,] -0.7142857  0.14285714  0.2857143
[2,]  1.0000000  0.00000000 -0.5000000
[3,] -0.0952381 -0.04761905  0.2380952
> solve(A) %*% A
     [,1]          [,2] [,3]
[1,]    1 -1.110223e-16    0
[2,]    0  1.000000e+00    0
[3,]    0  0.000000e+00    1
>
>
> A <- matrix(seq(from=1,to=9), nrow=3, byrow=T)
> det(A)
[1] 0
> solve(A)
Error in solve.default(A) :
  Lapack routine dgesv: system is exactly singular:
U[3,3] = 0
```

Cofactors

Lets review the cofactor way of looking at this (around page260 in the book). The cofactor way is a way of simplifying the ‘big formula’.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & & \\ & a_{22} & a_{23} \\ & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ a_{21} & & a_{23} \\ a_{31} & & a_{33} \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ a_{21} & a_{22} & \\ a_{31} & a_{32} & \end{vmatrix}$$

The cofactors along row 1 are $C_{1j} = (-1)^{1+j} \det M_{1j}$.

The cofactor expansion is $\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$. (11)

$$det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + a_{i3}C_{i3} \cdots + a_{in}C_{in}$$

Inverses: What are they

The matrix **A** is invertible IF there exists a matrix $\mathbf{A}^{-1}$ such that $\mathbf{A}\,\mathbf{A}^{-1} = \mathbf{I}$. This is similar to defining an inverse function, which swaps the domain and co-domain. Or think of it like the multiplicative inverse of numbers ($3 * \frac{1}{3} = 1$)

One important property is that if there is a nonzero vector x such that Ax is brought to 0 (i.e. has a non-trivial nullspace) then cannot have an inverse.

You can use solve(A) in R to find the inverse, since without the second parameter, the function assumes an identity matrix for B and returns the inverse of A.

Cofactor Method for Inverses

We know one method for finding the inverse: elimination on page83. Lets use co-factors as another method. We saw cofactors earlier when dealing with determinants. If we create a matrix C full of another matrix's co-factors, we can multiply A and C^T to get:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \times \begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix} = \begin{bmatrix} det(A) & 0 & 0 \\ 0 & det(A) & 0 \\ 0 & 0 & det(A) \end{bmatrix}$$

This allows us to see that the inverse of A is simply the cofactor matrix divided by the det(A): $\mathbf{A^{-1} = C^T / det(A)}$.

Eigenvectors and Eigenvalues

From the lecture notes: When you compute Ax, you are multiplying the matrix A onto vector x and this typically causes x to change its direction. However, there are some special vectors x which lie in the same direction as Ax. In other words, $Ax = \lambda x$ where λ is a scalar. If we are lucky enough to find such vectors, we call them eigenvectors and the corresponding λ as eigenvalues

If we think of a matrix as doing a linear transformation, eigen vectors point in the direction where the transformation is just a stretching, not a rotation.

Extra (from page 289):

- **The product of the n eigenvalues equals the determinant**
- The sum of the n eigenvalues equals the sum of the n diagonal entries.

Calculating Eigenvectors and values

```
Ax = ƛx
Ax - ƛx = 0
Ax - ƛIx = 0 (to make A and ƛ have same dimensions)
(A - ƛI)x = 0
```

So we are looking for non-trivial elements in the nullspace of A - ƛI. This is only the case for a non-invertible matrix. This only happens when:

```
det( A - ƛI ) = 0
```

Its from here we do our calculation, since we know how to find the det(). Page 287 has examples; so does lecture notes

Singular Value Decomposition

We have seen a ‘decomposition’ before. When we did elimination, we eventually found a LU ‘factorization’ or ‘decomposition’, where L and U were lower and upper triangular matrices. Singular Value Decomposition is another way of taking a matrix and splitting it into other matrices.

From the notes: “SVD is solving two related eigenvalue problems by constructing two square matrices from the original rectangular matrix and taking an eigen decomposition of the two matrices and relating them back to the original matrix”

SVD, cont

We can prove for a fact that every matrix that is m x n in shape can be broken down into

$A = U \Sigma V^T$

where U is an m × m orthonormal matrix, V is n × n orthonormal matrix, and Σ is m × n such it only has non-zero entries σ ii and all other entries are 0 when i != j. From this formula, we can find A^TA and AA^T which then shows us that, V and U are the eigenvectors of those terms respectively.

From link further down: " … for the intuition behind the SVD is that you can take any matrix (linear map) and break it up into three pieces: a rotation about the origin, a rescaling of each coordinate, followed by another rotation about the origin" -- a geometric approach

DS Look at SVD

https://jeremykun.com/2016/04/18/singular-value-decomposition-part-1-perspectives-on-linear-algebra/ -- this link empahsizes that the best thing about SVD isn't the factorization, but that it gives you an sorted list of best rank-k approximations to A, which can be used to approximate A but use less dimensions (think image compression).

Another link with an animation regarding PCA / SVD: http://stats.stackexchange.com/questions/2691/making-sense-of-principal-component-analysis-eigenvectors-eigenvalues/140579#140579

Latent Semantic Indexing

Extra: LSI is a method in information retrieval that uses SVD on a special matrix to help find hidden (ie latent) structure between words and documents. Main take away: