Where does linear algebra go from here?

The material from 1013 and 1014 includes a lot of topics. Among others, you learned about

1. vector spaces, subspaces, dimension
2. linear transformations and eigenvalues
3. orthogonal projection and inner products
4. Markov chains and dynamical systems

Today I'll offer a brief and informal sketch of how these show up in pure mathematics, internet search algorithms, psychology, and signals processing.

Fourier analysis

We've seen that $\mathbb{P}_n$, the polynomials of degree less than or equal to n, form a vector space of dimension $n+1$. Taking all polynomials together, we get an infinite dimensional vector space whose vectors are functions. However, we can be more general than this. Define $\mathcal{C}$ to be the set of functions which are integrable on the interval $[-\pi, \pi]$.

Fourier analysis

We've seen that $\mathbb{P}_n$, the polynomials of degree less than or equal to n, form a vector space of dimension $n + 1$. Taking all polynomials together, we get an infinite dimensional vector space whose vectors are functions. However, we can be more general than this. Define $\mathcal{C}$ to be the set of functions which are integrable on the interval $[-\pi, \pi]$.

We can define an *inner product* (remember, this is one name for a dot product) on this vector space:

$$f \cdot g = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx.$$

Fourier analysis

We've seen that $\mathbb{P}_n$, the polynomials of degree less than or equal to n, form a vector space of dimension $n + 1$. Taking all polynomials together, we get an infinite dimensional vector space whose vectors are functions. However, we can be more general than this. Define $\mathcal{C}$ to be the set of functions which are integrable on the interval $[-\pi, \pi]$.

We can define an *inner product* (remember, this is one name for a dot product) on this vector space:

$$f \cdot g = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx.$$

Then the functions

$$\mathcal{B} = \left\{ \frac{1}{\sqrt{2}}, \sin x, \cos x, \sin 2x \cos 2x, \sin 3x, \cos 3x, \dots \right\}$$

form a basis for $\mathcal{C}$.

This basis is orthonormal!

The Google PageRank algorithm

(exposition based on Higham and Taylor, *The Sleekest Link Algorithm*, 2003)

A search engine does 3 things:

1. Find web pages and store pertinent information in some sort of archive;
2. When queried, search archive to find a list of relevant pages;
3. Decide what order to display the found pages to the searcher.

Google's success with the third is a neat application of linear algbera.

PageRank

We'll model the internet as a collection of points, one for each webpage. Point A has an arrow connecting it to Point B if Page A has a link to Page B. We can record this information in an *adjacency matrix* that has $a_{ij} = 1$ exactly when page i links to page j.

We assume a page is important if

- lots of pages link to it, or
- important pages link to it.

PageRank

PageRank assigns a value r_j^n to the jth page at the nth iteration:

$$r_j^n = (1 - d) + d \sum_{i=1}^{N} \frac{a_{ij} r_i^{n-1}}{\deg_i}.$$

Here $0 < d < 1$ and $\deg_i$ is the number of arrows leaving page i.
Given some initial ranking (for every page, $1 - d$), the ranking of a page changes each time we iterate the ranking process. Note the following:

- endorsements from pages with high ranking increase the ranking
- endorsements from many pages increase the ranking
- each page gets the same initial influence, since we divide by the number of endorsements given out (the degree)

Iterating repeatedly, the numerical values assigned to each page stabilise over time, and Google will display the top ranked pages first to the searcher.

PageRank

I claimed that iterating the ranking leads to stable values for each page, but I didn't explain why. This argument is a bit more involved, but here are some of the ingredients:

- Given a system of linear equations described by $A\mathbf{x} = \mathbf{b}$, suppose we have a guess for $\mathbf{x}$.
- The *Jacobi iteration* is a process for turning our initial guess into a new guess, and PageRank turns out to be the Jacobi iteration applied to a system derived from the linking data.
- Under appropriate hyptheses, these guesses converge to an actual solution. (Compare this to Newton's Method for finding roots of a differentiable function.)
- This sort of technique comes from the field of *numerical linear algbera*;

Personality Tests and Dimension

Personality tests (e.g., Myers-Briggs, Big Five) classify an individual's personality in terms of a small number of traits (4 and 5, respectively).

Question

It's easy to list many, many traits that contribute to someone's personality, so why should a small list like this be interesting or useful?

Personality Tests and Dimension

Personality tests (e.g., Myers-Briggs, Big Five) classify an individual's personality in terms of a small number of traits (4 and 5, respectively).

Question
It's easy to list many, many traits that contribute to someone's personality, so why should a small list like this be interesting or useful?

Suppose everyone in the room listed traits that characterise personality, and let's say that we ended up with 100 different traits. Let's also assume that for each one of these, it's possible to assign each person a numerical score. Then each person could be assigned a point in $\mathbb{R}^{100}$.

Personality Tests and Dimension

When Myers and Briggs analysed the data for many, many people, they found that the corresponding points in $\mathbb{R}^{100}$ aren't randomly distributed, but in fact form a 4-dimensional subspace.

(Careful! I'm lying just a bit, but it's the right idea.)

The claim that personality is 4-dimensional is really a claim that there are four *independent* facets of personality, so determining these four determines the other 96.

Question

Given a vector space V and points sampled from a subspace W, how can you determine the dimension of W?