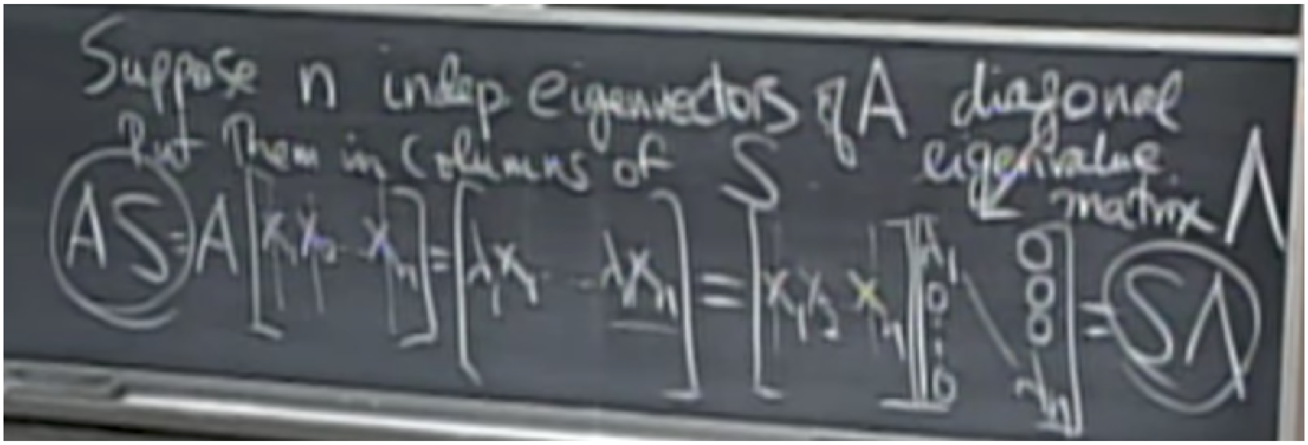
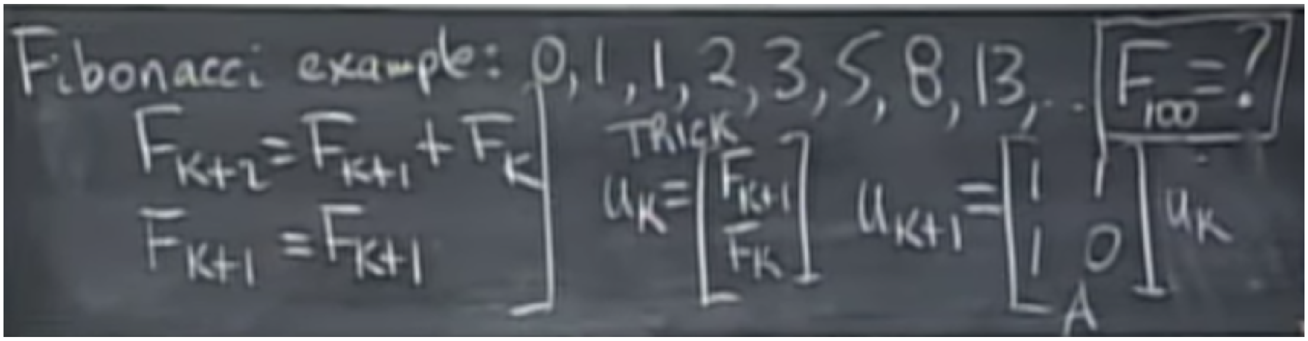
determinants, eigenvalues and eigenvectors

- a determinant function that maps a matrix to a number such that:
  - $|I| = 1$
  - exchanging rows reverses the sign of the determinant. (we can show that going from a matrix to another matrix can only be done with either an odd or even number of row exchanges → so this is well defined, otherwise the determinant wouldn't be a well defined function).
  - if we multiply the first row of A by t, then the determinant of that matrix is t times the determinant of A
  - if the first row of matrix A is $a + a', b + b'$ etc, then the determinant of A is the same as sum of determinants of A except with first row a, b etc + determinant of A except with first row a', b' etc.
- some consequences of the above definition:
  - the above implies that if two rows are equal, then the determinant is 0. (hint: interchange those rows, the determinant should change signs).
  - this also means that if a matrix isn't invertible, then the determinant is 0
  - if we do row k = row k - a times row i, then the determinant doesn't change (prove this, its easy).
  - if we have a row of 0's then the determinant is 0 (singular matrices have determinant 0)
  - if we have an upper triangular matrix, then the determinant is the product of the diagonal elements. (proof hint: start with the identity, and get to the upper triangular). $|A| = d_1..d_n$ when A is upper triangular.
  - We can use the above property to find that the formula for the 2-d determinant is $ad - bc$. (the well known formula), but derived from our basic definition of a determinant.
  - $|AB| = |A||B|$. proof: you can easily see why this is true for diagonal matrices. reducing a singular matrix down to its diagonals doesn't change the determinant as we have seen already. so to see that it is true for any invertible matrix, just reduce the matrix down to its diagonal form (which doesn't change the determinant) and we have already shown it for diagonal matrices, so hence proved. we can also prove this for the case where either A or B is not invertible.
  - it follows from the above that $|A^{-1}| = \frac{1}{|A|}$ and $|A^2| = |A|^2$.
  - also using the linearity property, $|2A| = 2^n|A|$ where n is the dimensions of the matrix.
  - $|A| = |A^T|$ (to see why reduce to row echelon form and tranpose it and compute the determinant of that matrix)
  - the consequence of the above property is that column operations also don't change the determinant (because a column operation is just a row operation of the transpose). also a 0 column means the determinant is 0. exchanging columns changes the sign. all the good stuff holds for columns too.
  - lets develop a formula for computing the determinant of a general matrix. first break up the given matrix as follows: break up the matrix into a sum of n matrices such that in each of the n matrices only 1 element in the first row is non-0 and the rest are all 0's. do similar breakdowns for the rest of the rows. notice that we now get n^n matrices with most of them having 0 determinant and only $n!$ have non-zero determinant. to compute the determinant of each of these $n!$ matrices and add them up, note that for some of them we need to do even number of column exchanges and for some an odd number. give appropriate signs for each. we find that determinant is the sum of determinants of

a_{11}*((det of the matrix obtained by removing 1st column, 1st row) - a_{12}*((det of the matrix obtained by removing 2nd column, 1st row, etc.
  - the term in the brackets above is called the cofactor. the cofactor of C_{ij} is the determinant of the (n-1) x (n-1) matrix formed by removing row i and column j from the n x n matrix, along with the appopriate sign.
  - given this, we can write the determinant as $a_{11}C_{11} + a_{12}C_{12}...a_{1n}C_{1n}$.
  - formula for matrix inverse: $A^{-1} = \frac{1}{det(A)}C^T$. To verify that this is true, we need to prove that $\frac{1}{det(A)}AC^T = I$. To see why this is true, see why the diagonal elements of AC^T will just be $det(A)$. All the non-diagonal elements will be of the form $a_{21}C_{11} + a_{22}C_{12} + a_{2n}C_{1n}$ where these will be 0's because this is like taking the determinant of a matrix with repeated rows (think about why). Hence the formula is proved. If its non-invertible (singular), then the determinant will be 0 and the inverse not defined.
  - Cramer's rule is the observation that if you have $A\vec{x} = \vec{b}$, then $x_k = \frac{det(B_k)}{det(A)}$ where B_k is basically A except with the k^{th} column replaced by $\vec{b}$. this is pretty easy to see if we think in terms of the cofactor matrix → no big deal here given the way we have approached determinants until now.
  - we can show that the determinant of a 3x3 matrix gives us the volume of a parallelo-box in 3-d whose vertices are defined by the 3 vectors that are the rows (or columns) of the matrix whose determinant we are computing.
- $Ax = \lambda x$. given a matrix there are some vectors which when put thru the "matrix function", the output is a vector which is in the same direction as the input vector. these are called the eigenvectors of the matrix A and the corresponding λ's are the eigenvalues that correspond to those eigenvectors.
- Notice that if A is singular, then $\lambda = 0$ is obviously an eigenvalue, since there are many things in the nullspace of A.
- One way to approach $Ax = \lambda x$ is to rewrite this as $(A - \lambda I)x = 0$. Now for values of λ where $A - \lambda I$ is singular, we will have interesting x's (eigenvectors) → in the nullspace of $A - \lambda I$. so we interested in finding λ's where $A - \lambda I$ is singular.
- we take $det(A - \lambda I) = 0$ and find the roots. this will give us the singular matrices. (this equation will have n roots, in general so we'll have as many eigenvalues, of course there will be repeats). this equation is called the characteristic equation.
- once we find the different λ's we will get a couple of singular matrices, and for each we find the basis for the null space and then we know how to get eigenvectors.
- Now, notice that once we find a matrix and its eigenvalues, if we add nI to the matrix, the corresponding eigenvalues are $\lambda + n$. proof: we have $Ax = \lambda x$ → $(A + nI)x = \lambda x + nx = (\lambda + n)x$. hence proved.
- Note the above is not true in general for $A + B$, only for $A + nI$.
- note that these determinants could very well turn out to have complex roots, in which case we have complex eigenvalues.
- note that if we have an upper triangular matrix, since the determinant is the product of the diagonal elements, we see that the eigenvalues are just the elements on the diagonal. (do $det(A - \lambda I) = 0$ to see why this is true).
- note the following matrix factorization where x_1, x_2 etc are the eigenvectors. We show below see pic) that $AS = S\Lambda$. If S (the matrix whose columns are eigenvectors) is invertible (has independent eigenvectors), then we can write $A = S\Lambda S^{-1}$



- Note that there are a small number of matrices that don't have independent eigenvectors and for those, we can't write $A = S\Lambda S^{-1}$
- this process is called diagonalization. its a eigenvector matrix times a diagonal matrix (the corresponding eigenvalues) times the inverse of the eigenvector matrix.
- We have seen A = LU factorization, A = QR (Gram Schmidt) and this is a third factorization (again, called diagonalization)
- Note that $tr(AB) = tr(BA)$. To verify this, see that $tr(AB) = \sum_i \sum_j a_{ij}b_{ji}$ and see that $tr(BA) = \sum_j \sum_i b_{ij}a_{ji}$. hence proved.
- what happens to eigenvalues if we square the matrix? $Ax = \lambda x$ → $A^2x = \lambda Ax = \lambda^2 x$ → squaring a matrix squared its eigenvalues too without changing its eigenvector! We can also verify this by $A^2 = S\Lambda S^{-1}S\Lambda S^{-1} = S\Lambda^2 S^{-1}$. we see the eigenvalue matrix is getting squared but the eigenvector doesn't change. In general, $A^k = S\Lambda^k S^{-1}$. This is powerful way to compute an arbitrarily large power of a matrix.
- Using the fact that $A = S\Lambda S^{-1}$ and that $tr(ABC) = tr(CAB)$, notice that the sum of eigenvalues = trace of matrix A and that the product of eigenvalues is equal to the determinant.
- Using the above, A^k → 0 as k → ∞ if all $|\lambda_i| < 1$. Nice! remember, all this assumes we have a full independent set of eigenvectors for all this to work. otherwise, we cannot diagonalize like this.
- A is sure to have n independent eigenvectors if all its n eigenvalues are different (proof). If you have repeated eigenvalues, you may or may not have independent eigenvectors.
- Look at this slick application of eigenvalues to find the fibonacci in a closed form:



Again, the trick is finding the eigenvalues and eigenvectors, diagonalize and then we are ready to find any arbitrary number in the fibonacci sequence.

- Notice that in the above example, the fib is just an example of a difference equation. Writing the initial values, and expressing the relation as a matrix, computing eigenvalues, we can find a good, fast way to compute any arbitrary value in the sequence. Sweet application!
- $\vec{x}^t = A\vec{x}$ is a system of first order linear differential equations where A is a n x n matrix and $\vec{x}$ is a vector whose components are the unknown functions. (solving the diff. eq system means finding $\vec{x}$). Lets see if $\vec{x}(t) = \vec{\eta}e^{rt}$ is a solution. differentiating both sides, $\vec{x}^t(t) = r\vec{\eta}e^{rt}$. If $\vec{x}(t) = \vec{\eta}e^{rt}$ is to be solution, it must satisfy $r\vec{\eta}e^{rt} = A\vec{x} = A\vec{\eta}e^{rt}$, so we must have $r\vec{\eta}e^{rt} = A\vec{\eta}e^{rt}$. Cancelling e^{rt} (because the exponentials cannot be 0, otherwise we couldn't cancel it), we ahve $A\vec{\eta} = r\vec{\eta}$. which means r would be an eigenvalue and η a corresponding eigenvector. so we must have $(A - rI)\vec{\eta} = \vec{0}$. Once we find the eigenvalues and eigenvectors, we can form solutions using $\vec{x}(t) = \vec{\eta}e^{rt}$.
- Professor Strang goes into some cases where we get complex eigenvalues that we don't discuss here.
- Notice that the eigenvalue of a particular solution determines the power of that e is raised to. so the solution function is stable if the eigenvalue λ < 0 and it blows up over time if λ > 0. Thus we see how the eigenvalue is critical to determining stability here.
- We have skipped over some details of diff eqs here but this is just an example application of eigenvalues/eigenvectors.
- A markov matrix (or stochastic matrix) is one in which every element is $0 \leq x \leq 1$ and the elements of a row add up to 1. you can see how this has applications in probability.
- a markov matrix times a markove matrix is a markov matrix (verify this)
- a markov matrix always has one of its eigenvalues 1. to see why, consider that that $A - I$ makes it singular because all the rows add up to 1. since $A - I$ is singular, $\lambda = 1$ is an eigenvalue.
- Every other eigenvalue is ≤ 1. to see why consider that $A^n x$ doesn't blow up with increasing n (because A^n is also a markov matrix). now $A^n x = \lambda^n x$, so it must be that $|\lambda_i| \leq 1$.
- Remember that $A^k = c_1\lambda_1^k x_1 + ...c_n\lambda_n^k x_n$. For high enough k only the term with $\lambda_i = 1$ can be used the approximate A^k. so $A^k \approx c_1\lambda_1^k x_1$ where λ_1 is the eigenvalue that has value 1. See how much the eigenvalue tells us now!
- A and A^T have the same eigenvalues. why? because if $A - rI$ is singlar, then $A^T - rI$ is also singular for that value of r. that's one reason why.
- Given an orthonormal basis $q_1..q_n$ it is easy to find $c_1, ...c_n$ such that $v = c_1q_1 + ...c_nq_n$. just do v^Tq_k to get c_k (check to see why this is true). This lets us express any vector in terms of an orthonormal basis easily.
- the fourier series is similar in idea to the above but is for function spaces. each function is an infinite dimensional vector (think of each value $g(x)$ as a dimension). orthogonality in vector space means $a^Tb = 0$. note that here we multiply each component and sum it up. the parallel of this in function space is $\int_{k_1}^{k_2} f(x)g(x)dx$ where the function is periodic with periodicity $k_2 - k_1$.
- we can break up $f(x)$ into a linear combination of basis components with each of the basis elements being orthogonal to one another so like $f(x) = a_0 + a_1cos(x) + b_1sin(x) + a_2cos(2x) + b_2sin(2x)...$. and we can find the coefficients by finding the dot product, called "norm". also we can show that the "inner prouducts" of the basis components are all 0.
- Strang just introduces the parallels in ideas between function spaces and vector spaces and the fourier series. He does not prove for example, that this set of basis can span any function (or even some subset of functions).