Lecture 30: Diagonalization

Diagonalization

Two matrices are called **similar** if $S^{-1}AS$. A matrix is called **diagonalizable** if it is similar to a diagonal matrix.

A matrix is diagonalizable if and only if it has an eigenbasis, a basis consisting of eigenvectors.

Proof. If we have an eigenbasis, we have a coordinate transformation matrix S which contains the eigenvectors v_i as column vectors. To see that the matrix $S^{-1}AS$ is diagonal, we check

$$S^{-1}ASe_i - S^{-1}Av_i = S^{-1}\lambda_i v_i = \lambda_i S^{-1}v_i = \lambda_i e_i \; .$$

On the other hand if A is diagonalizable, then we have a matrix S for which $S^{-1}AS = B$ is diagonal. The column vectors of S are eigenvectors because the k'th column of the equation $AS = BS$ shows $Av_i = \lambda_i v_i$.

Are all matrices diagonalizable? No! We need to have an eigenbasis and therefore that the geometric multiplicities all agree with the algebraic multiplicities. We have seen that the shear matrix

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

has the eigenvalues 1 for which the geometric multiplicity is smaller than the algebraic one. This matrix is not diagonalizable.

Simple spectrum

A matrix has simple spectrum, if all eigenvalues have algebraic multiplicity 1.

If a matrix has simple spectrum, then it is diagonalizable.

Proof. Because the algebraic multiplicity is 1 for each eigenvalue and the geometric multiplicity is always at least 1, we have an eigenvector for each eigenvalue and so n eigenvalues.

1 We have computed the eigenvalues of the rotation matrix

$$A = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

We have seen that the eigenvalues are $e^{i\alpha} = \cos(\alpha) + i\sin(\alpha)$, the eigenvectors are $\begin{bmatrix} \pm i \\ 1 \end{bmatrix}$. The eigenvectors are the same for every rotation-dilation matrix. With

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, S = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$$

we have

$$S^{-1}AS = \begin{bmatrix} a+ib & 0 \\ 0 & a-ib \end{bmatrix} \; .$$

Functional calculus

2 What is $A^{100} + A^{37} - 1$ if $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$? The matrix has the eigenvalues $\lambda_1 = 2 + \sqrt{3}$ with eigenvector $\vec{v}_1 = [\sqrt{3}, 1]$ and the eigenvalues $\lambda_2 = 2 - \sqrt{3}$. with eigenvector $\vec{v}_2 = [-\sqrt{3}, 1]$. Form $S = \begin{bmatrix} \sqrt{3} & -\sqrt{3} \\ 1 & 1 \end{bmatrix}$ and check $S^{-1}AS = D$ is diagonal. Because $B^k = S^{-1}A^kS$ can easily be computed, we know $A^{100} + A^{37} - 1 = S(B^{100} + B^{37} - 1)S^{-1}$.

Establishing similarity

3 Show that the matrices $A = \begin{bmatrix} 3 & 5 \\ 2 & 6 \end{bmatrix}$ $B = \begin{bmatrix} 4 & 4 \\ 3 & 5 \end{bmatrix}$ are similar. Proof. They have the same eigenvalues $8, 9$ as you can see by inspecting the sum of rows and the trace. Both matrices are therefore diagonalizable and similar to the matrix

$$\begin{bmatrix} 8 & 0 \\ 0 & 9 \end{bmatrix} \; .$$

- If A and B have the same characteristic polynomial and diagonalizable, then they are similar.
- If A and B have a different determinant or trace, they are not similar.
- If A has an eigenvalue which is not an eigenvalue of B, then they are not similar.
- If A and B have the same eigenvalues but different geometric multiplicities, then they are not similar.

Without proof we mention the following result which gives an if and only if result for similarity:

If A and B have the same eigenvalues with geometric multiplicities which agree and the same holds for all powers A^k and B^k, then A is similar to B.

Cayley Hamilton theorem

For any polynomial p,‘ we can form the matrix $p(A)$. For example, for $p(x) = x^2 + 2x + 3$, we have $p(A) = A^2 + 2A + 3$.

If f_A is the characteristic polynomial, we can form $f_A(A)$

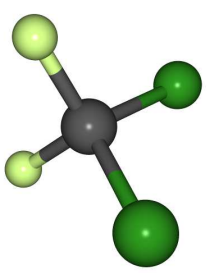
If A is diagonalizable, then $f_A(A) = 0$.

The matrix $B = S^{-1}AS$ has the eigenvalues in the diagonal. So $f_A(B)$, which contains $f_A(\lambda_i)$ in the diagonal is zero. From $f_A(B) = 0$ we get $Sf_A(B)S^{-1} = f_A(A) = 0$.

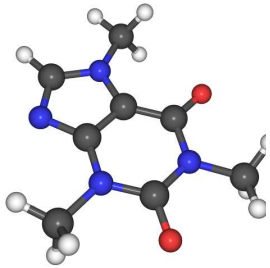
The theorem holds for all matrices: the coefficients of a general matrix can be changed a tiny bit so that all eigenvalues are different. For any such perturbations one has $f_A(A) = 0$. Because the coefficients of $f_A(A)$ depend continuously on A, they are zero in general.

An application in chemistry

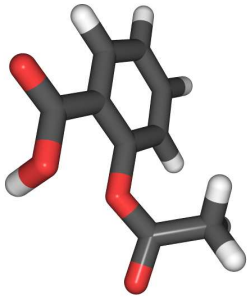
While quantum mechanics describes the motion of atoms in molecules, the vibrations can be described classically, when treating the atoms as "balls" connected with springs. Such approximations are necessary when dealing with large atoms, where quantum mechanical computations would be too costly. Examples of simple molecules are white phosphorus P_4, which has tetrahedronal shape or methan CH_4 the simplest organic compound or **freon**, CF_2Cl_2 which is used in refrigerants.



Freon CF_2Cl_2



Caffeine $C_8H_{10}N_4O_2$



Aspirin $C_9H_8O_4$

1

Let x_1, x_2, x_3, x_4 be the positions of the four phosphorus atoms (each of them is a 3-vector). The inter-atomar forces bonding the atoms is modeled by springs. The first atom feels a force $x_2 - x_1 + x_3 - x_1 + x_4 - x_1$ and is accelerated in the same amount. Let's just chose units so that the force is equal to the acceleration. Then

$$\begin{aligned} \ddot{x}_1 &= (x_2 - x_1) + (x_3 - x_1) + (x_4 - x_1) \\ \ddot{x}_2 &= (x_3 - x_2) + (x_4 - x_2) + (x_1 - x_2) \\ \ddot{x}_3 &= (x_4 - x_3) + (x_1 - x_3) + (x_2 - x_3) \\ \ddot{x}_4 &= (x_1 - x_4) + (x_2 - x_4) + (x_3 - x_4) \end{aligned}$$

which has the form $\ddot{x} = Ax$, where the 4×4 matrix

$$A = \begin{bmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{bmatrix} \quad v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, v_4 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$ are

the eigenvectors to the eigenvalues $\lambda_1 = 0, \lambda_2 = -4, \lambda_3 = -4, \lambda_3 = -4$. With $S = [v_1 v_2 v_3 v_4]$, the matrix $B = S^{-1}BS$ is diagonal with entries $0, -4, -4, -4$. The coordinates $y_i = Sx_i$ satisfy $\ddot{y}_1 = 0, \ddot{y}_2 = -4y_2, \ddot{y}_3 = -4y_3, \ddot{y}_4 = -4y_4$ which we can solve y_0 which is the center of mass satisfies $y_0 = a + bt$ (move molecule with constant speed). The motions $y_i = a_i \cos(2t) + b_i \sin(2t)$ of the other eigenvectors are oscillations, called **normal modes**. The general motion of the molecule is a superposition of these modes.

Homework due April 20, 2011

1 What is the probability that an upper triangular 3×3 matrix with entries 0 and 1 is diagonalizable?

2 Which of the following matrices are similar?

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

3 Diagonalize the following matrix in the complex:

$$A = \begin{bmatrix} 2 & -3 & 0 & 0 \\ 3 & 2 & 0 & 0 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 6 & 5 \end{bmatrix}$$

[1] We grabbed the pdb Molecule files from *http : //www.sci.ouc.bc.ca*, translated them with "povchem" from .pdb to .pov rendered them under Povray.