Symmetric Matrices 4/13/2001	Math 21b, O. Knill

Homework for Monday April 16: Section 7.3, Numbers 2,10,12,16,26*,36

SYMMETRIC MATRICES appear everywhere. For example in **geometry** as **generalized dot products** $v \cdot Av$ in **statistics** as **correlation matrices** $\mathrm{Cov}[X_k, X_l]$, in quantum mechanics as **observables**, in **neural networks** as **learning maps** $x \mapsto \mathrm{sign}(Wx)$, in graph theory as **adjacency matrices** etc. etc. The reason, why symmetric matrices play a special role, is that they play amoung matrices the role of real numbers among the complex numbers. Their eigenvalues (which often have physical or geometrical interpretations) are real. Furthermore, one can calculate with symmetric matrices like for example find to a given matrix A a matrix B such that $B^2 = A$ (one calls such a B square root of A.) Try to find a matrix B such that $B^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$...

RECALL. We have seen when an eigenbasis exists, a matrix A can be transformed into a diagonal matrix $B = S^{-1}AS$, where $S = [v_1, ..., v_n]$. The matrices A and B are called **similar** and B is called the **diagonalisation** of A.

SIMILAR MATRICES. Similar matrices have the same characteristic polynomial $\det(S^{-1})(A - \lambda)S) = \det(A - \lambda)$ and have therefore the same determimant, trace and eigenvalues. Physisics, who call the set of eigenvalues **the spectrum** say, that they are isospectral. The spectrum is what you "see" (ethymologically the name comes from the fact that in quantum mechanics the spectrum of radiation gets associated with eigenvalues of matrices.)

QUESTION. How large is the set of matrices which can be diagonalized? We have seen that matrices with disjoint eigenvalues can be diagonalized. Symmetric matrices form an important class. They appear often in applications.

SPECTRAL THEOREM. Symmetric matrices A can be diagonalized $B = S^{-1}AS$ with an orthogonal S. The diagonal entries of B are real and the eigenvalues of A.

PROOF. If all eigenvalues are different, we have an eigenbasis and we can diagonalize. In that case, the eigenvectors are orthogonal because $\lambda_i v_i \cdot v_j = Av_i \cdot v_j = v_i \cdot A^T v_j = v_i \cdot Av_j = \lambda_j v_i \cdot v_j$ is only possible if $v_i \cdot v_j = 0$. Therefore $B = S^{-1}AS$ is diagonal and real because A and S are real. In general, call B the diagonal matrix with eigenvalues as entries, we can change the matrix A to $A = A + (C - A)t$ where C is a matrix with pairwise different eigenvalues. Then the eigenvalues are different for most t different from 0. (If $\lambda_i(t) - \lambda_j(t)$ were zero for t on some interval, then we had $\lambda_i(t) = \lambda_j(t)$ for all t which contradicts in the case $t = 1$.) The orthogonal matrix S_t converges for $t \to 0$ coefficientwise to an orthogonal matrix S.

EXAMPLE 1. The matrix $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$ has the eigenvalues $a + b, a - b$ and the eigenvectors $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} /\sqrt{2}$ and $v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} /\sqrt{2}$. You see that they are orthogonal. The orthogonal matrix $S = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$ diagonalizes A.

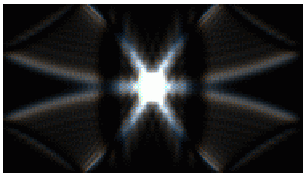
EXAMPLE 2. (see homework) The $n \times n$ matrix $A = \begin{bmatrix} 1 & \cdots & 1 \\ \cdots & \cdots & \cdots \\ 1 & \cdots & 1 \end{bmatrix}$ has $(n-1)$ eigenvalues 0 to the eigenvectors $\begin{bmatrix} ? & \cdots & ? \cdots & ? \end{bmatrix}$ and one eigenvalue n to the eigenvector $\begin{bmatrix} ? & \cdots & ? \cdots & ? \end{bmatrix}$. All these vectors can be made orthogonal and a diagonalisation is possible even so the eigenvalues have multiplicities.

EXAMPLE 3. The symmetric $n \times n$ matrix $A = \begin{bmatrix} 0 & 1 & 0 & \cdot & 0 & 1 \\ 1 & 0 & 1 & \cdot & \cdot & 0 \\ 0 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & 0 \\ 0 & \cdot & \cdot & 1 & 0 & 1 \\ 1 & 0 & \cdot & 0 & 1 & 0 \end{bmatrix}$ has the complex eigenvectors $v_j = \begin{bmatrix} e^{2\pi i 1 j/n} & e^{2\pi i 2j/n} & \cdots & e^{2\pi i n j/n} \end{bmatrix}$. The real and imaginary parts, when normalized form real orthogonal eigenvectors to the eigenvalues $\lambda_j = 2\cos(2\pi j/n)$. A linear independent subset forms an orthonormal basis. This diagonalisation is the **discrete Fourier transform**.

SQUARE ROOT. How do we find a square root of a given symmetric matrix? Because $S^{-1}AS = B$ is diagonal and we know how to take a square root of the diagonal matrix B, we can form $C = S\sqrt{B}S^{-1}$ which satisfies $C^2 = S\sqrt{B}S^{-1}S\sqrt{B}S^{-1} = SBS^{-1} = A$.

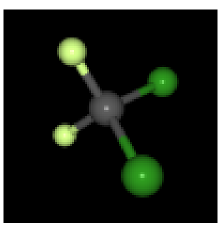
EXHIBITION. "Where do symmetric matrices occur?" The next boxes are motivation.

I) PHYSICS: A **quantum mechanical** system is described with a vector $v(t)$ which depends on time. The evolution is given by the **Schroedinger equation** $\dot{v} = i\hbar Lv$, where L is a symmetric matrix and $\hbar$ is a small number called the Planck constant. As for any linear differential equation, one has $v(t) = e^{i\hbar Lt}v(0)$. If $v(0)$ is an eigenvector to the eigenvalue λ, then $v(t) = e^{it\hbar\lambda}v(0)$. Physical observables are given by symmetric matrices too, L representing the energy. If the system is represented by $v(t)$, the value of the observable $A(t)$ is $v(t) \cdot Av(t)$. For example, if v is an eigenvector to an eigenvalue λ of the energy matrix L, then the energy of $v(t)$ is λ.
Sometimes, it is good to switch the perspective and let evolve the observables instead of the vector v.



This is called the **Heisenberg picture**. In order that $v \cdot A(t)v = v(t) \cdot Av(t) = S(t)v \cdot AS(t)v$ we have $A(t) = S(T)^*AS(t)$, where $S^* = \overline{S^T}$ is the correct generalization of the adjoint to complex matrices. $S(t)$ satisfies $S(t)^*S(t) = 1$ which is called **unitary** and the complex analogue of orthogonal. The matrix $A(t) = S(t)^*AS(t)$ has the same eigenvalues as A and is **similar** to A.

II) CHEMISTRY. **Adjecency matrix** of a graph. A graph with n vertices which can be connected through edges is determined by the adjecency matrix A_{ij} which is 1 if the two vertices i, j are connected and zero otherwise. The matrix A is symmetric. The eigenvalues λ_j are real and can be used to analyze the graph. One interesting question is: how many graphs do have the same eigenvalues then a given graph.
In chemistry, one is interested in that because it allows to make rough computations of the electron density distribution in molecules. In this **Hückel theory**, the molecule is represented as a graph. The eigenvalues λ_j of that graph represent the energies the molecule can have, and the eigenvectors describe the electron density distribution.



In the **freon molecule** for example (which me met last time already), there are 5 atoms and the adjacency matrix is
$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} .$$
This matrix has the eigenvalue 0 with multiplicity 3 (the three dimensional kernel is obtained immediatly from the fact that 4 rows are the same) and the eigenvalues ± 2. The eigenvector to the eigenvalue ± 2 is $\begin{bmatrix} \pm 2 & 1 & 1 & 1 & 1 \end{bmatrix}^T$.

III) MECHANICS. The **Toda lattice** is a particle system on the line, where neighboring particles attract each other with the force e^{-d}, where d is the distance. If q_n are the positions of the particles, then the Newton equations are $\frac{d^2}{dt^2}q_n = e^{q_{n+1}-q_n} - e^{q_n - q_{n-1}}$. A coordinate transformation $a_n^2 = e^{q_{n+1}-q}, 2b_n = p_n = \dot{q}_n$ brings the system into the form $\dot{a}_n = a_n(b_{n+1} - b_n)$, $\dot{b}_n = 2a_{n-1}^2 - 2a_n^2$. It can be written in matrix form as $\dot{A} = [B, A] = BA - AB$, where
$$A = \begin{bmatrix} b_1 & a_1 & 0 & \cdot & 0 & a_n \\ a_1 & b_2 & a_2 & \cdot & \cdot & 0 \\ 0 & a_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_{n-2} & 0 \\ 0 & \cdot & \cdot & a_{n-2} & b_{n-1} & a_{n-1} \\ a_n & 0 & \cdot & 0 & a_{n-1} & b_n \end{bmatrix}, \; B = \begin{bmatrix} 0 & a_1 & 0 & \cdot & 0 & -a_n \\ -a_1 & 0 & a_2 & \cdot & \cdot & 0 \\ 0 & -a_2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & a_{n-2} & 0 \\ 0 & \cdot & \cdot & -a_{n-2} & 0 & a_{n-1} \\ a_n & 0 & \cdot & 0 & -a_{n-1} & 0 \end{bmatrix}.$$
With $\dot{S} = BS$, one has $A(t) = S^{-t}A(0)S^t$. The matrix $A(t)$ is similar to A. Because the eigenvalues of A are preserved, one has enough conserved quantities which allow to find closed-form solutions of these differential equations. Evenso nonlinear, it can be linearized. The continuum version of this system is called the **KdE equation** and a model for water waves.

IV) STATISTICS. If we have a random vector $X = [X_1, \cdots, X_n]$ and $E[X_k]$ denotes the expected value of X_k, then $[A]_{kl} = E[(X_k - E[X_k])(X_l - E[X_l])]$ is called the **covariance matrix** of the random vector X. It is a symmetric matrix. Diagonalizing this matrix $B = S^{-1}AS$ produces new random variables which are **uncorreleated**.