**SYMMETRIC MATRICES** 

**Math 21b, O. Knill**

**SYMMETRIC MATRICES.** A matrix $A$ with real entries is symmetric, if $A^T = A$.

**EXAMPLES.** $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ is symmetric, $A = \begin{pmatrix} 1 & -1 \\ -1 & 3 \end{pmatrix}$ is not symmetric.

**EIGENVALUES OF SYMMETRIC MATRICES.** Symmetric matrices $A$ have real eigenvalues.

**PROOF.** The dot product is extend to complex vectors as $(v, w) = \sum \bar{v}_i w_i$, for real vectors it satisfies $(v, w) = \bar{v} \cdot w$ and has the property $(Av, w) = \bar{\lambda} (v, w)$ for real matrices $A$ and $(\lambda v, w) = \bar{\lambda} (v, w)$ as well as $(v, Av) = \lambda (v, v)$ and $(\lambda v, \lambda w) = \lambda^2 (v, w)$.

**EXAMPLE.** $A = \begin{pmatrix} p & -q \\ q & p \end{pmatrix}$ has eigenvalues $\pm \sqrt{p^2 + q^2}$ which are real if and only if $q = 0$.

**EIGENVECTORS OF SYMMETRIC MATRICES.** Symmetric matrices $A$ have an orthonormal eigenbasis.

**PROOF.** If $Av = \lambda v$ and $Aw = \mu w$. The relation $\lambda (v, w) = (Av, w) = (Av, w) = (\lambda w, w)$ is only possible if $v = 0$ or $\lambda = \mu$.

**RECALL.** We have seen when an eigenbasis exists, a matrix $A$ can be transformed to a diagonal matrix $B = S^{-1}AS$, where $S = [v_1, ..., v_n]$. The matrices $A$ and $B$ are similar. $B$ is called the diagonalization of $A$. Similar matrices have the same characteristic polynomial $det (B - \lambda I) = det (S^{-1}(A - \lambda I)S) = det (A - \lambda I)$ and have therefore the same determinant, trace and eigenvalues. Physicists call the set of eigenvalues also the **spectrum**. They say that these matrices are isospectral. The spectrum is what you “see” (etymologically the same origin comes from the fact that in quantum mechanics the spectrum of radiation can be associated with eigenvalues of matrices.)

**SPECTRAL THEOREM**

Symmetric matrices $A$ can be diagonalized $B = S^{-1}AS$ with an orthogonal $S$.

If all eigenvalues are different, there is an eigenbasis and diagonalization is possible. The eigenvectors are all orthogonal and $B = S^{-1}AS$ is diagonal containing the eigenvalues. In general, 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 all except finitely many $t$. The orthogonal matrices $S_t$ converges for $t \rightarrow 0$ to an orthogonal matrix $A$ and $S_t$ diagonalizes $A$.

**EXAMPLE 1.** The matrix $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ has the eigenvalues $\lambda = a \pm b$ and the eigenvectors $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. They are orthogonal. The orthogonal matrix $S = [v_1 \ v_2]$ diagonalizes $A$.

**EXAMPLE 2.** The $3 \times 3$ matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$ has 2 eigenvalues 0 to the eigenvectors $[1 \ -1 \ 0]$ and one eigenvalue 3 to the eigenvector $[1 \ 1 \ 1]$. All these vectors can be made orthogonal and a diagonalization is possible even so the eigenvalues have multiplicities.

**SQUARE ROOT OF A MATRIX.** 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 $C^2 = S\sqrt{B}S^{-1}S\sqrt{B}S^{-1} = SBS^{-1} = A$.

**RAYLEIGH FORMULA.** We write also $(\bar{v}, \bar{w}) = \bar{v} \cdot \bar{w}$. If $E(t)$ is an eigenvector of length 1 to the eigenvalue $\lambda(t)$ of a symmetric matrix $A(t)$ which depends on $t$, differentiation of $\lambda(t) = \bar{\lambda}(t)\bar{E}(t)$ with respect to $t$ gives $(A(t) - \lambda(t)E(t))E(t) = 0$ with respect to $t$ gives $(A(t) - \lambda(t)E(t)) = 0$. The symmetry of $A - \lambda I$ implies $0 = (\bar{v}, (A - \lambda I)(\bar{w} - \lambda \bar{w}))$. We see that the Rayleigh quotient $\lambda = \bar{v}^T(A\bar{v})$ is a polynomial in $t$ if $A(t)$ only involves terms $t^2, t^3, ...$ only. The formula shows how $\lambda$ changes, when $t$ changes. For example, $\lambda(t) = \begin{pmatrix} 1 \\ t^2 \\ 1 \end{pmatrix}$ for $t = 2$ the eigenvector $\bar{v} = [1, 1] / \sqrt{2}$ to the eigenvalue $\lambda = 5$. The formula tells that $\lambda(t^2) = (A(t)/t^2 \bar{E}(t) \bar{E}) = \begin{pmatrix} 0 & 4 \\ 0 & 0 \end{pmatrix}$, $\bar{v} \bar{w} = 4$. Indeed, $\lambda(t) = 1 + t^2$ has at $t = 2$ the derivative $2t = 4$.

**EXHIBITION.** Where do symmetric matrices occur? Some informal motivation:

**I) PHYSICS:** In quantum mechanics a system is described with a vector $\psi(t)$ which depends on time $t$. The evolution is given by the Schroedinger equation $i\psi = i\hbar \lambda(t)$, where $\lambda(t)$ is a symmetric matrix and $\hbar$ is a small number called the Planck constant. As for any linear differential equation, one has $\psi(t) = e^{i\lambda(t)\hbar}V(0)$. If $V(0)$ is an eigenvector to the eigenvalue $\lambda$, then $\psi(t) = e^{i\lambda(t)\hbar}V(0)$. Physical observables are given by symmetric matrices too. The matrix $\lambda(t)$ represents the energy. Given $\psi(t)$, the value of the observable $\lambda(t)$ is $\psi(t)$. For example, if $\psi$ is an eigenvector to an eigenvalue $\lambda$ of the energy matrix $\lambda(t)$, then the energy of $\psi(t) = \lambda$. This is called the Heisenberg picture. In order to $\lambda(t) = \psi(t) \cdot \lambda(t) \psi(t) \cdot \lambda(t) \psi(t)$ we have $\lambda(t) = \lambda(t)$ if $\lambda(t)$ 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 $\lambda(t) = S(t)^*AS(t)$ has the same eigenvalues as $A(t)$ and is similar to $A$.

**II) STATISTICS.** If we have a random vector $X = [X_1, ..., X_n]$ and $E[X_i]$ denotes the expected value of $X_i$, then $\{A_{ij} = E[X_i - E[X_i][X_j - E[X_j]]] = E[X_iX_j - E[X_i]E[X_j]]\}$ is called the covariance matrix of the random vector $X$. It is a symmetric $n \times n$ matrix. Diagonalizing this matrix $B = S^{-1}AS$ produces new random variables which are uncorrelated.

For example, if $X$ is the sum of two dice and $Y$ is the value of the second dice then $E[X] = \begin{pmatrix} 3.5 \\ 3.5 \end{pmatrix}$ and $E[Y] = \begin{pmatrix} 3.5 \\ 3.5 \end{pmatrix}$. The matrix entry $A_{11} = E[X^2] - E[X]^2 = \begin{pmatrix} 7 \end{pmatrix}$. The matrix entry $A_{12} = E[XY] - E[X]E[Y] = \begin{pmatrix} 7 \end{pmatrix}$. The covariance matrix is the symmetric matrix $A = \begin{pmatrix} 35/6 & 35/12 \\ 35/12 & 35/12 \end{pmatrix}$.