NOTES FOR 19 NOV (TUESDAY)

1. Recap

(1) Proved the spectral theorem for self-adjoint operators.
(2) As consequences, discussed the Rayleigh quotient, operator norms, and the singular value decomposition.

2. Normal operators; Spectral theorem for normal operators

Can an orthonormal basis of eigenvectors be found for more general operators ? Indeed. Here is a necessary condition : A normal operator $T : V \to V$ between finite-dimensional inner product spaces is an operator such that $TT^* = T^*T$ (likewise, a normal matrix). If an operator has an orthonormal basis of eigenvectors, it is normal : Indeed, if we choose any orthonormal basis and write everything using matrices, $UTU^\dagger = D$ and hence, $UT^\dagger U^\dagger = D^\dagger$ and $UTT^\dagger U^\dagger = DD^\dagger = UT^\dagger TU^\dagger$. Thus, $TT^\dagger = T^\dagger T$.

Here is an important spectral theorem.

Theorem 2.1. *An operator $T : V \to V$ between finite-dimensional inner product spaces has an orthonormal basis of eigenvectors iff T is normal.*

To prove the harder direction, we first notice the following.

Lemma 2.2. *Let $T : V \to V$ be a normal operator between finite-dimensional inner product spaces. Then eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Proof. Since T is normal, $\|(T - cI)v\| = \|(T^* - \bar{c}I)v\|$ and hence $Tv = cv$ iff $T^*v = \bar{c}v$. Therefore, if $Te_i = \lambda_i e_i$, then $(Te_i, e_j) = \lambda_i(e_i, e_j)$. Hence, $\lambda_j(e_i, e_j) = (e_i, T^*e_j) = \lambda_i(e_i, e_j)$. Thus, $(e_i, e_j) = 0$. □

Actually, we prove something more general : If e is an eigenvector of T, and W is the orthogonal complement of the space generated by e, then W is an invariant subspace for T and T^*. Indeed, if $(w, e) = 0$, then $(Tw, e) = (w, T^*e) = \lambda(w, e) = 0$ (and likewise for T^*). Using this observation, we can prove the spectral theorem exactly as before. Indeed, if e is an eigenvector of the normal of T, then since the restriction $T : W \to W$ is also normal (why ?) inductively W has an orthonormal basis of eigenvectors. That basis along with e forms a basis of V.

There is another proof of this result (which is important in its own right). Firstly,

Lemma 2.3. *Let A be the matrix of $T : V \to V$ in an orthonormal basis such that A is upper triangular. Then T is normal iff A is diagonal.*

Proof. If A is diagonal, we are trivially done. If T is normal, then $Te_1 = A_{11}e_1$. Hence, $T^*e_1 = \bar{A}_{11}e_1$. But, $(T^*e_1) = \sum_j \bar{A}_{1j}e_j$. Hence, $A_{1j} = 0$ for all $j > 1$. Thus, $Te_2 = A_{22}e_2$ and so on (actually, more elegantly, we are done by induction). □

Now we have the Schur decomposition lemma.

Lemma 2.4. *Let $T : V \to V$ be any operator over a finite-dimensional inner product space. Then there is an orthonormal basis of V in which T is upper-triangular.*

Proof. We induct on the dimension on V. For $n = 1$ it is trivial. Assume truth for $1, 2 \ldots, n-1$. Choose a unit eigenvector e of T. Let W be the orthogonal complement of the space generated by e. Now consider $\Pi_W \circ T : W \to W$. By the induction hypothesis, there is an orthonormal basis of W such that $\Pi_W \circ T$ is upper-triangular. Adjoining e to that basis, we conclude the result. □

As a corollary, we see that if T is normal, then in the above orthonormal basis, T is diagonal.

3. Bilinear forms

Let V be a finite-dimensional real vector space and $T : V \times V \to \mathbb{R}$ be a symmetric bilinear form (in this case, called a real quadratic form). We have an important theorem (Sylvester's law of inertia) : If V is a real vector space, then there exists an ordered basis e_i such that $T(e_i, e_j) = \delta_{ij} p_i$ where $p_i = 0$ or 1 or -1. The number of $1, 0, -1$ is determined uniquely by T. (The signature of a real quadratic form is defined as $n_+ - n_-$. The rank of a quadratic form is $n_+ + n_-$.)

Proof. Firstly, if e_i is any basis, let $B_{ij} = T(e_j, e_i) = T(e_i, e_j) = B_{ji}$. Then $T(v, w) = w_i B_{ij} v_j = w^T B v$, where B is a symmetric matrix. It has real eigenvalues. Note that if we change the basis to $e_{old,i} = P_{ji} e_{new,j}$, then $T(v, w) = w_{old}^T P^T B P v_{old}$.Change the basis using an orthogonal matrix O such that $O^T B O = D$. Now we can further rescale the basis vectors so that the positive eigenvalues become 1 and the negative ones -1. Let V_+ be the span of the eigenvectors with positive eigenvalues (dimension n_+), and V_- (dimension n_-) that of negative eigenvalues. Thus, T is negative-definite on V_- and positive-definite on V_+.

As for uniqueness, let V_0 be the subspace of V such that if $v \in V_0$, then $T(v, w) = 0$ for all $w \in V$ (and thus $V = V_0 \oplus V_+ \oplus V_-$) which means that $w^T B v = 0$ for all w. Thus, $Bv = 0$. Hence, V_0 is the kernel of B (and so p_0 is uniquely determined). Let W be any subspace on which T is positive-definite. We claim that W, V_0, V_- are independent. Indeed, if $c_1 w + c_2 v_0 + c_3 v_- = 0$, then $0 = c_1 T(w, w) + c_3 T(v_-, w)$. Likewise, $c_3 T(v_-, v_-) + c_1 T(w, v_-) = 0$. Thus, $0 \leq c_1^2 T(w, w) = c_3^2 T(v_-, v_-) \leq 0$. Hence W, V_0, V_- are independent. Hence, $dim(W) \leq n_+$. So, if we "diagonalise" the quadratic form in two different ways, comparing n_+ for both, we see that they are equal. Likewise for $n-$. □

The classification of conic sections can be done using this theorem. Indeed, if $z^2 = x_1^2 + x_2^2 + \ldots$ and $cz + a_1 x_1 + a_2 x_2 + \ldots = 0$, then we can write it as $v^T B v = k$ where k is some constant. Using Sylvester's law of inertia, we see that one can further change the variables linearly to conclude that $\sum_i p_i v_i^2 = k$ where p_i are $0, 1$, or -1. Hence, in principle, we can know all possible conic sections (note however that the form of B is somewhat restricted).