NOTES FOR 14 NOV (THURSDAY)

1. Recap

(1) Defined self-adjoint operators and gave examples.
(2) Proved that any two norms on f.d. spaces are equivalent, defined the operator norm, and showed that $I-A$ is invertible if $\|A\|_{op} < 1$, and that e^A makes sense. Defined unitary operators and proved that they correspond to unitary matrices in an orthonormal basis.

2. Unitary operators; Spectral theorem for self-adjoint operators

In whatever follows, as usual we assume that the spaces are finite-dimensional. (Spectral theory, that is, the theory of eigenvalues and eigenvectors is much more subtle in infinite-dimensions.)

Theorem 2.1. *The eigenvalues of a self-adjoint operator are real. Moreover, eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Proof. If $Av = \lambda v$, then $(v, Av) = \bar{\lambda}(v, v)$ and $(Av, v) = \lambda(v, v)$. Since $A = A^*$, $\lambda = \bar{\lambda}$.
If $Aw = \mu w$, then $\lambda(v, w) = (Av, w) = (v, Aw) = \bar{\mu}(v, w) = \mu(v, w)$. Hence, $(v, w) = 0$. □

As a corollary, if a self-adjoint operator has distinct eigenvalues, it has an orthonormal basis of eigenvectors, i.e., for a Hermitian matrix A there exists a unitary matrix U such that $UAU^\dagger = D$. (Such a matrix is said to be "unitarily equivalent" to a diagonal matrix.) In fact, much more can be said.

Theorem 2.2. *Let $T : V \to V$ be a self-adjoint operator between finite-dimensional inner product spaces. Then, there exist invariant subspaces E_i and orthogonal projections $\Pi_i : V \to E_i$ such that $I = \sum_i \Pi_i$ and $T : E_i \to E_i$ is of the form $Tv = \lambda_i v$ where $\lambda_i \in \mathbb{R}$. In particular, T is diagonalisable using an orthonormal basis.*

Proof. Induct on $n = dim(V)$. For $dim(V) = 1$ it is trivial. Assume it has been proven for $1, \ldots, n-1$. Now there exists a unit eigenvector e of T with real eigenvalue λ. Let Π_e be the orthogonal projection to the space $< e >$ spanned by e. Then $I = \Pi_e + (I - \Pi_e)$. We claim that $W =< e >^\perp$ is an invariant subspace. Indeed, if $(v, e) = 0$, then $(Tv, e) = (v, Te) = (v, \lambda e) = \lambda(v, e) = 0$. Moreover, the restriction $T : W \to W$ is self-adjoint (why?) Hence, by the induction hypothesis, $\sum_i \Pi_i = I$ on W. Thus, $v = \Pi_e v + (I - \Pi_e)v = \Pi_e v + \sum_i \Pi_i(I - \Pi_e)v$. Note that $\Pi_i(I - \Pi_e)v = \Pi_i v - \Pi_i \Pi_e v = \Pi_i v - \Pi_i(v, e)e = \Pi_i v$. If $\lambda = \lambda_i$ for some i, define a new projection operator as $\Pi_i + \Pi_e$. We are done. □

As a consequence, if A is a Hermitian matrix, then there exists an orthonormal basis w_i such that $A = \sum_i \lambda_i |w_i\rangle\langle w_i|$.
Here are a couple of consequences.

(1) We have a variational characterisation of the largest eigenvalue of a Hermitian matrix : Let A be a self-adjoint operator over a finite-dimensional inner product space and let λ be its largest eigenvalue. Then $\lambda = \sup_{\|v\|=1}(Av, v)$. Moreover, $\|A\| = \sup |\lambda|$ over all eigenvalues of A.
Indeed, let e_i be an eigenvector basis. Then, $(Av, v) = \sum_i \lambda_i |v_i|^2 \leq \lambda \|v\|^2 = \lambda$ with equality holding iff v is an eigenvector corresponding to λ.

(2) As a HW exercise, you will show that simultaneous diagonalisation with orthonormal eigenvectors is possible for families of commuting self-adjoint operators. Now we have the following result (functional calculus).

Theorem 2.3. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function. There exists a unique function $\tilde{f} : Herm_{n\times n}(\mathbb{C}) \to Herm_{n\times n}(\mathbb{C})$ such that $\tilde{f}(A)v = f(\lambda)v$ for every eigenvector v of A.*

Indeed, since $v = \sum_i (v, e_i)e_i$, $\tilde{f}(A)v = \sum_i (v, e_i) f(\lambda_i)e_i$, i.e., $\tilde{f}(A) = U^\dagger f(D)U$. (Actually, the theorem is stronger : It says that if f is continuous, then so is $\tilde{f}$.) As a consequence,

Proposition 2.4. *Let A be any matrix. Then $C = A^\dagger A$ has a unique Hermitian non-negative definite square root B, i.e., $B^2 = A^\dagger A$.*

Proof. Firstly, a Hermitian matrix is non-negative definite iff its eigenvalues are non-negative. Indeed, if $Av = \lambda v$, then $(v, Av) = \lambda(v, v) \geq 0$. Moreover, if its eigenvalues are non-negative, we get the result by diagonalisation.
Secondly, C is clearly Hermitian and non-negative definite and hence $\sqrt{C}$ exists as $B = U^\dagger \sqrt{D} U$. In fact, choosing the positive square root of D, clearly, B is Hermitian and non-negative definite. Indeed, $v^\dagger B v = (Uv)^\dagger \sqrt{D} Uv \geq 0$. Noticing that the eigenvalues of B^2 are simply the eigenvalues of B squared (by simultaneously diagonalising them), we are done with uniqueness too. □

(3) The operator norm of a Hermitian matrix is $\max_\lambda |\lambda|$ where the maximum is over all eigenvalues of A. Indeed, $\|Av\| = \|\sum_i \lambda_i v_i e_i\|$ where e_i is an orthonormal basis of eigenvectors. So, $\|Av\| \leq |\lambda|_{max}\|v\|$ with equality holding iff v is an eigenvector.

(4) In fact, one can compute the norm of even non-rectangular matrices $A \in Mat_{m\times n}$ using eigenvalues in a clever manner. Indeed, since $A^\dagger A$ and $AA^\dagger$ are Hermitian (and non-negative definite), they are diagonalisable (with eigenvalues being non-negative numbers). Let $A^\dagger A = UD_1U^\dagger$ (the columns of U are called the "left-singular" vectors of A) and $AA^\dagger = VD_2V^\dagger$ (likewise, right-singular vectors).
Claim : The non-zero entries of D_1 and D_2 are the same.
Proof : If $A^\dagger Av = \lambda v$, then $AA^\dagger Av = \lambda Av$. Hence, if $\lambda \neq 0$ since $Av \neq 0$, Av is an eigenvector of $AA^\dagger$ with eigenvalue λ. Moreover, if $v_1, v_2, \ldots, v_k$ are linearly independent eigenvectors of $A^\dagger A$ with eigenvalue λ, then Av_i are also linearly independent. Indeed, if $\sum_i c_i Av_i = 0$, then $\sum_i c_i A^\dagger Av_i = \sum_i c_i \lambda v_i = 0$ and hence $c_i = 0 \ \forall \ i$. Likewise, the same argument shows that if v is an eigenvector of $AA^\dagger$, then $A^\dagger v$ is one of $A^\dagger A$. Thus, the dimensions of their non-zero eigenspaces are equal.
Now, $\|Av\|^2 = (Av, Av) = (A^\dagger Av, v)$ and hence the operator norm of Av is the square root of the largest eigenvalue of $A^\dagger A$. In the HW you will show that there are unitary matrices U, V such that $V^\dagger AU$ is "diagonal" and consists of zeroes as well as the square roots of the non-zero eigenvalues of $A^\dagger A$. These numbers are called the singular values of A. (This decomposition is called the singular value decomposition (SVD).)