2015 Chicago Math REU Apprentice Program Exercises
Friday, July 24

Instructor: László Babai
Scribe: Angela Wu

Theorem 25.1. If $D \subseteq \mathbb{R}^n$ is compact (closed and bounded) and non-empty and $f : D \to \mathbb{R}$ is continuous, then $f$ achieves an argmax on $D$.

Exercise 25.2. Recall that for $A \in M_n(\mathbb{R})$ with $A = A^T$, the Rayleigh quotient is defined as $R_A(x) = \frac{x^T A x}{x^T x}$. For symmetric $\phi : V \to V$, the Rayleigh quotient is defined as $R_\phi(x) = \frac{\langle x, \phi x \rangle}{||x||^2}$. Use the above theorem to show that argmax $x R_\phi(x)$ is achieved on $V \setminus \{0\}$.

Proof. Note that $R_\phi(\alpha x) = R_\phi(x)$. So any value achieved by $R_\phi$ on $V \setminus \{0\}$ is achieved on the unit sphere $\{x \in V | ||x|| = 1\}$. Restrict $x$ from $V \setminus \{0\}$ to the unit sphere. The unit sphere is compact, so the maximum is achieved. \hfill \Box

25.1 Continuation of Hoffman-Singleton

Exercise 25.3 (Exercise 24.30). Let $f(x) = x^4 + ax^3 + bx^2 + cx - 15$, with $a, b, c \in \mathbb{Z}$. If $p \in \mathbb{Q}$ and $f(p) = 0$, then $p \in \mathbb{Z}$ and $p | 15$.

Proof. Suppose first that $p \in \mathbb{Z}$ and $f(p) = 0$. We claim that $p | 15$. But, then $p(p^3 + ap^2 + bp + c) = 15$, so $p | 15$.

Now, suppose that $f(p/q) = 0$ with $\gcd(p, q) = 1$. Then, expanding $f(p/q)$ and multiplying through by $q^4$, we arrive at $p^4 + ap^2q + bp^2q^2 + cpq^3 - 15q^4$. But then $q | p^4$, so $q = 1$. \hfill \Box

Exercise 25.4. If $f(x) = a_0 + a_1 x + \ldots + a_n x^n \in \mathbb{Z}[x]$ with $a_0a_n \neq 0$, then if $p/q \in \mathbb{Q}$ is a root with $\gcd(p, q) = 1$, then $p | a_0$ and $q | a_n$.

Theorem 25.5 (Hoffman-Singleton Theorem). Let $G$ be a $k$-regular graph with girth $\geq 5$ and $n = k^2 + 1$ vertices. Then $k \in \{1, 2, 3, 7, 57\}$.

Proof (continued). Recall: for the adjacency matrix $A$ of a regular graph of degree $k$, the all-ones vector $\vec{1}$ is an eigenvector with eigenvalue $k$. Let us normalize this vector: $e_1 := \frac{1}{\sqrt{n}} \vec{1}$. Let us extend $e_1$ to an orthonormal eigenbasis $e_1 \ldots e_n$ of $A$. So $e_i \perp e_1$ for all $i \geq 2$ and therefore $Je_i = 0$ ($i \geq 2$). Furthermore, we have seen that $A^2 + A - (k - 1)I = J$ and therefore, for $i \geq 2$ we have $(A^2 + A - (k - 1)I)e_i = 0$, so the eigenvalue $\lambda_i$ corresponding to $e_i$ satisfies $\lambda_i^2 + \lambda_i - (k - 1) = 0$. So the eigenvalues $\lambda_2, \ldots, \lambda_n$ take only two values:
\( \lambda_i = \frac{-1 + \sqrt{4k - 3}}{2} = -\frac{1}{2} \pm \frac{s}{2} \) with \( s = \sqrt{4k - 3} \). Changing notation, let \( \lambda_0 = k, \lambda_1 = (-1 + s)/2, \) and \( \lambda_2 = (-1 - s)/2 \). The multiplicity of \( \lambda_0 = k \) is 1, and let \( m_1 \) be the multiplicity of \( \lambda_1 \) and \( m_2 \) the multiplicity of \( \lambda_2 \). To solve for \( m_1 \) and \( m_2 \), notice that:

\[
1 + m_1 + m_2 = n = k^2 + 1.
\]

\[
0 = \text{Tr}(A) = k + m_1 \lambda_1 + m_2 \lambda_2.
\]

Substituting the values of the \( \lambda_i \) and the equation \( m_1 + m_2 = k^2 \) in the second equation, we obtain \( 2k - k^2 + s(m_1 - m_2) = 0 \). So, either (1) \( m_1 = m_2 \) or (2) \( 4k - 3 \) is a square. In case (1) we have \( k = 2 \). In case (2), substituting \( k = (s^2 + 3)/4 \) in the equation \( k^2 - 2k - s(m_1 - m_2) = 0 \) we obtain \( s^4 - 2s^2 - 16s(m_1 - m_2) - 15 = 0 \). Thus, \( s = \pm 1, \pm 3, \pm 5, \pm 15 \), so \( k = 1, 3, 7, 57 \).

### 25.2 Operator norm

**Definition 25.6.** Let \( B \in \mathbb{R}^{k \times n} \). The operator norm of \( B \) is

\[
\|B\| = \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|Bx||}{|x|} = \max_{|x|=1} |Bx||/|x||.
\]

In particular, \( (\forall x)(|Bx| \leq \|B\| \cdot |x|) \).

**Theorem 25.7.** If \( A \in M_n(\mathbb{R}), A = A^T \), then \( \|A\| = \max_i |\lambda_i| \).

**Proof.** Let \( \tau = \max_i |\lambda_i| = |\lambda_j| \).

(1) \( \|A\| \geq \tau \). Let \( x = e_j \) be the eigenvector to \( \lambda_j \), so \( |Ae_j| = \tau |e_j| \).

(2) \( \|A\| \leq \tau \). Use the Spectral Theorem. Let \( \xi = (e_1 \ldots e_n) \) be an ONB and let \( A = [\phi]_{\xi} \).

Pick an orthonormal eigenbasis \( b = (b_1 \ldots b_n) \). Let \( x = \sum \alpha_i b_i \), \( \phi(x) = \sum \alpha_i \phi(b_i) = \sum \alpha_i \lambda_i b_i \).

\[
|x| = (\sum \alpha_i^2)^{1/2} \text{ and } |Ax| = (\sum \alpha_i^2 \lambda_i)^{1/2} \leq (\sum \alpha_i^2)^{1/2} = (\tau^2 \sum \alpha_i^2)^{1/2} = \tau |x| .
\]

**Exercise 25.8.** Let \( A \) be symmetric. Then, \( A \) is positive semidefinite \( \iff \) \( (\forall i)(\lambda_i \geq 0) \).

**Proof.** If \( \phi \) is positive semidefinite, then \( (\forall x)(\langle x, \phi x \rangle \geq 0) \). Write \( x = \sum \alpha_i b_i \), so \( \langle x, \phi x \rangle = \langle \sum \alpha_i b_i, \sum \alpha_i \lambda_i b_i \rangle = \sum \lambda_i \alpha_i^2 \geq 0 \). Since this needs to hold for all values of the \( \alpha_i \), we must have \( \lambda_i \geq 0 \) for all \( i \).

**Exercise 25.9.** Prove: \( B^T B \) is symmetric and positive semidefinite.

**Theorem 25.10.** Let \( B \in \mathbb{R}^{k \times n} \). Then, \( \|B\| = \sqrt{\lambda_{\max}(B^T B)} \).

**Proof.** \( B^T B \) is symmetric and positive semidefinite. We calculate:

\[
\frac{\|Bx\|^2}{|x|^2} = \frac{x^T (B^T B) x}{|x|^2} = R_{B^T B}(x),
\]

and this quantity is maximized at \( \lambda_{\max}(B^T B) \).

**Exercise 25.11.** Let \( B = (b_{ij}) \). Prove: for every \( i, j \) we have \( \|B\| \geq |b_{ij}| \).
25.3 Card Shuffling: Finite Markov Chains, Rapid Mixing

Let $G$ be a $d$-regular graph. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of the adjacency matrix $A$ for $G$.

Exercise 25.12. Show that (1) $\lambda_1 = d$, and (2) $(\forall i)(|\lambda_i| \leq d)$.

Exercise 25.13. Let $G$ be a regular graph. Show that $\lambda_2 = \lambda_1$ if and only if $G$ is disconnected.

Exercise 25.14. Show: if $G$ is bipartite then $\lambda_n = -\lambda_1$, $\lambda_{n-1} = -\lambda_2$, etc. (The spectrum is symmetric about the origin.)

Definition 25.15. The eigenvalue gap is $d - \rho$ where $\rho = \max_{i \geq 2} |\lambda_i|$.

Remark 25.16. For regular graphs, the gap governs the speed of convergence of the random walk. By the preceding two exercises, for disconnected regular graphs and for (regular) bipartite graphs, the eigenvalue gap is 0 and indeed in these cases the random walk does not converge to the uniform distribution.

Definition 25.17. A finite Markov Chain (random walk) is given by a set of states, $1 \ldots n$, and a transition matrix $T = (p_{ij})_{n \times n}$. The MC is a stochastic process (sequence of random variables) $X_t$ ($t \in \mathbb{Z}, t \geq 0$); we think of $X_t$ as the location of a particle at time $t$ (or the state of a system at time $t$). The evolution of the process is governed by the transition probabilities $p_{ij} = \mathbb{P}[X_{t+1} = j \mid X_t = i]$.

Exercise 25.18 (Evolution of Markov Chain). Let $p_{ij}^{(k)} = \mathbb{P}[X_{t+k} = j : X_t = i]$ be the transition probability from state $i$ to state $j$ after $k$-steps. Show that $(p_{ij}^{(k)}) = T^k$.

Exercise 25.19. In general, the transition matrix is not symmetric and may have complex eigenvalues. Prove: 1 is always an eigenvalue of the transition matrix, and for all other eigenvalues $\lambda_i$ we have $|\lambda_i| \leq 1$.

Example 25.20. Consider the naive random walk on a $d$-regular graph $G$. The adjacency matrix $A = A_G$ is symmetric, with eigenvalues $\lambda_1 = d \geq \lambda_2 \geq \ldots \geq \lambda_n$. The transition matrix is described by $T = \frac{1}{d}A$ with eigenvalues $1 \geq \lambda_2/d \geq \ldots \geq \lambda_n/d$. Recall that $\rho = \max_{i \geq 2} |\lambda_i|$.

Theorem 25.21 (Convergence rate vs. eigenvalue gap). Consider Example 25.20. Then $(\forall i, j) \left( |p_{ij}^{(k)} - \frac{1}{n} | < \left( \frac{\rho}{d} \right)^k \right)$.

Proof. This will follow from Theorem 25.22 below in view of Ex. 25.11.

Theorem 25.22. $||T^k - \frac{1}{n}J|| \leq \left( \frac{\rho}{d} \right)^k$.

Proof. By Theorem 25.7, we need to show that the largest absolute value of eigenvalues of $T^k - \frac{1}{n}J$ is $\leq \left( \frac{\rho}{d} \right)^k$. 

3
Pick an orthonormal eigenbasis $e_1 \ldots e_n$ for $A$, with $e_1 = \frac{1}{\sqrt{n}} \vec{1}$ and $Je_1 = 0$ for $i \geq 2$. These are also eigenvectors for $J$: $Je_1 = ne_1$ and $Je_i = 0e_i$ for $i \geq 2$. Thus, $e_1 \ldots e_n$ is an eigenbasis for both $A$ and $J$. It follows that it is also an eigenbasis for $T^k - (1/n)J$. We use this to calculate the eigenvalues of $T^k - (1/n)J$ and thus find its norm.

Calculate:

$$
\begin{align*}
\left( T^k - \frac{1}{n}J \right) e_1 &= T^k e_1 - \frac{1}{n}Je_1 = e_1 - \frac{1}{n}ne_1 = 0. \\
\left( T^k - \frac{1}{n}J \right) e_i &= T^k e_i - \frac{1}{n}Je_i = \left( \frac{\lambda_i}{d} \right)^k e_i \quad (i \geq 2).
\end{align*}
$$

Thus the eigenvalues of $T^k - \frac{1}{n}J$ are $0, (\frac{\lambda_2}{d})^k, \ldots (\frac{\lambda_r}{d})^k$. So, $||T^k - \frac{1}{n}J|| = \max_{i \geq 2} |\frac{\lambda_i}{d}|^k = \left( \frac{\rho}{d} \right)^k$.

25.4 Singular Value Decomposition

**Problem:** Given $B \in \mathbb{R}^{k \times n}$, find $C \in \mathbb{R}^{k \times n}$ with $\text{rank}(C) \leq r$ that minimizes $||B - C||$.

The key to this problem is the Singular Value Decomposition, an extension of the Spectral Theorem to rectangular matrices.

**Theorem 25.23** (Spectral Theorem). For any $A \in M_n(\mathbb{R})$ with $A = A^T$, there exists $S \in O(n)$ such that $S^{-1}AS = D$ is diagonal.

**Remark 25.24.** The following are equivalent

- $S$ is an orthogonal matrix ($S \in O(n)$).
- The columns of $S$ are an orthonormal basis of $\mathbb{R}^n$.
- The rows of $S$ are an orthonormal basis of $\mathbb{R}^n$.
- $S^{-1} = S^T$.

**Theorem 25.25** (Singular Value Decomposition). Let $B \in \mathbb{R}^{k \times n}$. There exists $S \in O(k)$ and $T \in O(n)$ such that $S^{-1}BT = (d_{ij})$ satisfies $d_{ij} = 0$ as long as $i \neq j$ and $d_{11} \geq d_{22} \geq \ldots d_{mm} \geq 0$, where $m = \min\{k,n\}$. The nonzero diagonal entries are called the singular values $\sigma_i = d_{ii} > 0$ for $i = 1, \ldots, r$ ($r = \text{rank}(B)$) and the transformation $B \mapsto S^{-1}BT$ is called a singular value decomposition (SVD) of $B$.

**Proof.** The proof is a few lines using the Spectral Theorem. 

**Exercise 25.26.** If $B \in \mathbb{R}^{k \times n}$, $S \in O(k)$ and $T \in O(n)$, then $||B|| = ||S^{-1}BT||$.

**Exercise 25.27.** $\text{rank}(B^T)B) = \text{rank}(B)$.

**Theorem 25.28.** If the nonzero eigenvalues of $B^TB$ are $\lambda_1 \geq \ldots \geq \lambda_r > 0$, then the singular values of $B$ are $\sigma_i = \sqrt{\lambda_i}$, for $i = 1 \ldots r$.

**Proof.** Consider the SVD of $B$: $\Sigma = S^{-1}BT$, with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$. Then, $\Sigma^T \Sigma = (S^{-1}BT)^T(S^{-1}BT) = T^{-1}(B^T)BT$ has on the diagonal $\sigma_1^2, \sigma_2^2, \ldots$. 

Exercise 25.29 (Transpose of a linear map). Let $V$ and $W$ be Euclidean spaces and $\phi : V \to W$. Prove: there is a unique $\psi : W \to V$ such that for all $x \in V$ and $y \in W$ we have $\langle \phi x, y \rangle_W = \langle x, \psi y \rangle_V$. (Hint: next exercise)

This $\psi$ is called the transpose of $\phi$, denote $\phi^T$.

Exercise 25.30. Prove: if $g$ is an ONB in $V$ and $f$ and ONB in $W$ then $[\phi^T]_{g,f} = [\phi]_{e, \tilde{f}}$.

Exercise 25.31. Prove that $\phi : V \to V$ is (a) symmetric iff $\phi^T = \phi$; (b) orthogonal iff $\phi^T = \phi^{-1}$.

Exercise 25.32 (SVD theorem for linear maps). An equivalent form of the SVD theorem to transformations:

Let $\phi : V \to W$ have rank $r$. Then there exist numbers $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and an ONB $e$ of $V$ and an ONB $f$ of $W$ such that $\phi e_i = \sigma_i f_i$ and $\phi^T f_i = \sigma_i e_i$ for $i = 1, \ldots, r$.

Proof (from the SVD theorem for matrices). Let $V$ be an $n$-dimensional Euclidean space with ONB $e'$ and $W$ a $k$-dimensional Euclidean space with ONB $f'$. Let $\phi : V \to W$ and $B = [\phi]_{e', f'}$. So $S^{-1}BT = \Sigma$ as in the SVD theorem. Let us change the bases $e', f'$ via the basis change transformations corresponding to $S$ and $T$, respectively; call the new bases $e$ and $f$. Then $[\phi]_{\text{old}} = B$ and $[\phi]_{\text{new}} = \Sigma$. Then, $\phi e_i = \sigma_i f_i$ for $i = 1, \ldots, r$ by definition. The relation $\phi^T f_i = \sigma_i e_i$ follows by taking the transpose of the equation $S^TBT = \Sigma$.

Theorem 25.33. The closest rank-$s$ approximation $B^{(s)}$ of $B$ is found by erasing the smallest $r-s$ singular values of $\Sigma$ and transforming the resulting matrix $\Sigma^{(s)}$ back: $B^{(s)} = S \Sigma^{(s)} T^{-1}$.

Exercise 25.34. $||\Sigma - \Sigma^{(s)}|| = \sigma_{s+1}$, where $\Sigma^{(s)}$ is $\Sigma$ truncated to rank $s$ by replacing each $\sigma_j$, $j \geq s+1$, by zero (as in Theorem 25.33).

Proof of Theorem 25.33. Since multiplication by orthogonal matrices on either side preserves the norm, we have that $||B - B^{(s)}|| = \sigma_{s+1}$ by Exercise 25.34.

Claim. For any $C$ such that $\text{rank}(C) \leq s$ we have $||B - C|| \geq \sigma_{s+1}$.

Proof of the Claim. As in the proof of Ex. 25.32, let $V$ be an $n$-dimensional Euclidean space with ONB $e'$ and $W$ a $k$-dimensional Euclidean space with ONB $f'$. Let $\phi$ be the map corresponding to $B$, i.e., $B = [\phi]_{e', f'}$. Let moreover $\psi$ be the map corresponding to $C$, i.e., $C = [\psi]_{e', f'}$. Note that $\text{rank}(\psi) = \text{rank}(C) \leq s$. By the Rank-Nullity Theorem, $\dim(\ker(\psi)) = n - \dim(\text{im}(\psi)) \geq n - s$, where $n = \dim V$.

We need to show that $||\phi - \psi|| \geq \sigma_{s+1}$, i.e., there exists $x \in V$, $x \neq 0$ such that $||(\phi - \psi)x|| \geq \sigma_{s+1}||x||$.

Let now $\epsilon$ and $f$ be the pair of bases with respect to which $[\phi]_{e, f} = \Sigma$ as in Exercise 25.32. Let $U = \text{span}(e_1, \ldots, e_{s+1})$. So $\dim U = s+1$ and therefore $U \cap \ker(\psi) \neq \emptyset$. Let $x \in U \cap \ker(\psi)$, $x \neq 0$. We claim that $||(\phi - \psi)x|| \geq \sigma_{s+1}||x||$.

Indeed, $\psi x = 0$, so $(\phi - \psi)x = \phi x$. We can write $x = \sum_{i=1}^{s+1} \alpha_i e_i$ (because $x \in U$). So $||x|| = \left(\sum_{i=1}^{s+1} \alpha_i^2\right)^{1/2}$ and $\phi x = \sum_{i=1}^{s+1} \alpha_i \sigma_i f_i$. It follows that $||\phi x|| = \left(\sum_{i=1}^{s+1} \alpha_i^2 \sigma_i^2\right)^{1/2} \geq \sigma_{s+1} \left(\sum_{i=1}^{s+1} \alpha_i^2\right)^{1/2} = \sigma_{s+1}||x||$. This completes the proof of the Claim and thereby the proof of Theorem 25.33. \qed
Hermitian Spaces, Complex Version of Spectral Theorem

Definition 25.35. The adjoint (or conjugate transpose) of the complex matrix $A$ is the matrix $A^* = \bar{A}^T$. So if $A = (\alpha_{ij})$ then the $(i,j)$ entry of $A^*$ is $\bar{\alpha}_{ji}$. The map $(x,y) \mapsto x^*y$ is the standard Hermitian dot product on $\mathbb{C}^n$.

Exercise 25.36. If $x \in \mathbb{C}^n$ then $x^*x$ is real and positive unless $x = 0$.

Proof. $x^*x = \sum \bar{\alpha}_i \alpha_i = \sum |\alpha_i|^2 > 0$, unless $x = 0$.

Definition 25.37. $\|x\| = \sqrt{x^*x}$ (norm). The vectors $x, y \in \mathbb{C}^n$ are orthogonal if $x^*y = 0$. A matrix $A \in M_n(\mathbb{C})$ is Hermitian if $A^* = A$. A matrix $C \in M_n(\mathbb{C})$ is unitary if $C^{-1} = C^*$.

A matrix $B \in M_n(\mathbb{C})$ is normal if $A^*A = AA^*$.

So a real Hermitian matrix is a symmetric matrix, and a real unitary matrix is an orthogonal matrix.

Definition 25.38 (Complex Hermitian space). Let $V$ be a vector space over $\mathbb{C}$ with a positive definite sesquilinear inner product, i.e., and inner product that satisfies the same axioms as the inner product in real Euclidean space except $\langle \alpha x, y \rangle = \bar{\alpha} \langle x, y \rangle$ (but $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle$). Moreover, $\langle y, x \rangle = \overline{\langle x, y \rangle}$.

An example is $\mathbb{C}^n$ with the standard Hermitian dot product.


Definition 25.40. We say that the matrices $A, B \in M_n(\mathbb{C})$ are unitarily similar if there exists a unitary matrix $C$ such that $B = C^{-1}AC$.

Theorem 25.41. Every $A \in M_n(\mathbb{C})$ is unitarily similar to a triangular matrix.

Proof. Let $\mathcal{E}$ be an ONB of an $n$-dimensional Hermitian space $V$. Let $\phi : V \to V$ be defined by $[\phi]_{\mathcal{E}} = A$. We have already proved that $A$ is similar to a triangular matrix; this means there exists a basis $\mathcal{E}'$ such that $[\phi]_{\mathcal{E}'}$ is a triangular matrix. We observed that this is equivalent to saying that each subspace $U_i := \text{span}(e'_1, \ldots, e'_i)$ is $\phi$-invariant. Now orthogonalize $\mathcal{E}'$; this does not change the $U_i$. Then normalize this orthogonal basis; this again does not change the $U_i$. But we now have an ONB with respect to which $[\phi]$ is triangular, which is equivalent to the theorem.

Theorem 25.42. A matrix $A \in M_n(\mathbb{C})$ is unitarily similar to a diagonal matrix if and only if $A$ is normal.

Proof. $\Rightarrow$ is immediate from Ex. 25.43 below and the observation that every diagonal matrix is normal.

$\Leftarrow$: Use Theorem 25.41 and Exercises 25.43 and 25.44 (below).

Exercise 25.43. If $A$ is normal and $B$ is unitarily similar to $A$, then $B$ is normal.

Exercise 25.44. If an upper triangular matrix $T$ is normal then $T$ is diagonal.
Exercise 25.45.

(a) $A^* = A$ (A is Hermitian) $\iff$ A is normal and all the eigenvalues of $A$ are real

(b) $A^* = A^{-1}$ (A is unitary) $\iff$ A is normal and all the eigenvalues of $A$ have unit absolute value ($|\lambda_i| = 1$)

Enjoy your summer! (And feel free to send email to the instructor.)