Week 5; Constrained Optimization, SVD... Math 2240, Spring '24

1. The Rayleigh Quotient

Given a *square* matrix A, define $R_A : \mathbb{R}^n \to \mathbb{R}$ by

$$R_A(x) = \frac{\|Ax\|^2}{\|x\|^2}.$$

A surprising fact is that R_A takes its critical values at the eigenvalues of A, and that the critical values are precisely the eigenvalues of A.

1. Argue that if $x \neq 0$ is a critical point of R_A, then so is λx for any $\lambda \in \mathbb{R}$, and that R_A is constant on lines through the origin: $R_A(\lambda x) = R_A(x)$
2. Before finding the extrema, argue topologically that R_A attains a maximum and minimum on $\mathbb{R}^n$.
3. Restrict your attention to $R_A : S^{n-1} \to \mathbb{R}$. Find $x^* = \text{argmax}_{x \in S^{n-1}} R_A(x)$.
4. Show that x^* is the eigenvector of A corresponding to the largest eigenvalue, and that this eigenvalue is $max_{x \in \mathbb{R}^n} R_A(x)$.

2. SVD, but "by hand"

Let $A : \mathbb{R}^m \to \mathbb{R}^n$ a matrix, and $A^T : \mathbb{R}^n \to \mathbb{R}^m$ its transpose. We know that if A were square, we could apply the spectral theorem to get an orthonormal basis organized into a matrix U such that $A = U^T \Lambda U$ for Λ diagonal. We're hoping to extend this result.

- Show $AA^T : \mathbb{R}^n \to \mathbb{R}^n$ and $A^T A : \mathbb{R}^m \to \mathbb{R}^m$ are symmetric.
- Using spectral theorem, choose an orthonormal basis for $\mathbb{R}^m$ on which $A^T A$ is diagonal (with possible zeros). Can the eigenvalues be negative?
- Find an orthonormal basis for $ker(A)^\perp = im(A^T)$ such that $\mathbb{R}^m = im(A^T) \oplus ker(A) = A \oplus B$ (hint: show that $ker(A^T A) = ker(A)$).
- Find an orthonormal basis for $\mathbb{R}^n = im(A) \oplus ker(A^T) = C \oplus D$ such that $A : B \to D$ and $A^T : D \to B$ are each the zero map, and $A : A \to B$ is invertible. (hint: if $A^T A q_i = s_i^2 q_i$, do the vectors Aq_i form a nice basis for $\mathbb{R}^n$? What about $ker(A^T)$?)

The s_i determined above are called the *singular values* of A, the q_i are *right singular vectors*, and p_i are *left singular vectors*.

3. Another Rayleigh Quotient

The Rayleigh quotient can be generalized to non-square matrices. If $A : \mathbb{R}^m \to \mathbb{R}^n$, then now define $R_A : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by

$$R_A(y, x) = \frac{y \cdot Ax}{\|y\|\|x\|}$$

- Argue that, if we want to find the x, y which maximize R_A, it suffices to consider $R_A : S^{n-1} \times S^{m-1} \to \mathbb{R}$, and that from this we already know R_A attains a maximum and minimum.
- Express the manifold $S^{n-1} \times S^{m-1}$ as a regular value of some function $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^2$.
- Show that $\max_{x \in S^{n-1}, y \in S^{m-1}} R(y, x)$ is attained by the singular values of A. Are the (x^*, y^*) at which these are obtained the vectors we saw in (3.)? Are they uniquely determined?

4. Another Minimization Problem

Consider a linear system $Ax = v$ where $A \in \mathbb{R}^{n \times m}$. This has a solution if and only if $v \in \text{im}(A)$, which could be the case if A isn't onto. What we'd like is an *approximate* solution, $Ax = \hat{v}$, where $\hat{v}$ is the unique vector in $\text{im}(A)$ that is closest to v, i.e., the unique solution to the problem

$$\min_{\hat{v} \in \text{im}(A)} \|v - \hat{v}\|^2 = \min_{x \in \text{dom}(A)} \|Ax - v\|^2$$

Such a problem is called a *least-squares* problem, or a *minimum norm* problem when $\| \cdot \|$ is defined differently from $\|x\|^2 = \sum_i x_i^2$.

- Suppose we found $\hat{v} \in \text{im}(A)$ which minimizes the above norm. Show that if $x, y \in \text{dom}(A)$ solve the above problem, then $x - y \in \ker(A)$.
- Argue that the set of all least squares solutions is $x + \ker(A)$, where x is a particular solution.
- Of physical importance is the $y \in x + \ker(A)$ for which $\|y\|^2$ is smallest. Show that if y is the least-squares solution in $\text{dom}(A)$, then it is the unique vector in $\text{dom}(A)$ with this property
- Remember that $v - \hat{v} \perp \text{im}(A)$ by the projection property: the nearest vector $\hat{v}$ to v is that which has orthogonal error. Using this, show that $(A^T A)^{-1} A^T v \in \text{dom}(A)$ is the minimum norm least-squares solution.

5. Prep for Integration: Infinite sums?

For an arbitrary set $\mathcal{X} \subseteq \mathbb{R}^+$, define the sum $\sum_{x\in\mathcal{X}} x$ by

$$\sum_{x\in\mathcal{X}} x = \sup\{\sum_{x\in A} x \ : \ A \subseteq \mathcal{X} \ \text{ where } |A| \text{ finite}\}$$

Show that if $\mathcal{X}$ is uncountable, then $\sum_{x\in\mathcal{X}} x = \infty$. (hint: consider the sets $A_n = \{x \in \mathcal{X} : x > 1/n\}$)

So defining integration by "counting points in $\mathcal{X}$" is doomed to fail — there's too much to count!