Low Rank Approximation and Extremal Gain Problems

These notes pull together some similar results that depend on partial or truncated SVD or eigenvector expansions.

1 Low rank approximation

In lecture 15 we considered the following problem. We are given a matrix $A \in \mathbf{R}^{m\times n}$ with rank r, and we want to find the nearest matrix $\hat{A} \in \mathbf{R}^{m\times n}$ with rank p (with $p \leq r$), where 'nearest' is measured by the matrix norm, *i.e.*, $\|A - \hat{A}\|$. We found that a solution is

$$\hat{A} = \sum_{i=1}^{p} \sigma_i u_i v_i^T,$$

where

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

is the SVD of A. The matrix $\hat{A}$ need not be the only rank p matrix that is closest to A; there can be other matrices, also of rank p, that satisfy $\|A - \tilde{A}\| = \|A - \hat{A}\| = \sigma_{p+1}$.

It turns out that the same matrix $\hat{A}$ is also the nearest rank p matrix to A, as measured in the Frobenius norm, *i.e.*,

$$\|A - \hat{A}\|_F = \left(\mathbf{Tr}(A - \hat{A})^T(A - \hat{A})\right)^{1/2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}(A_{ij} - \hat{A}_{ij})^2\right)^{1/2}.$$

(The Frobenius norm is just the Euclidean norm of the matrix, written out as a long column vector.) In this case, however, $\hat{A}$ is the unique rank p closest matrix to A, as measured in the Frobenius norm.

2 Nearest positive semidefinite matrix

Suppose that $A = A^T \in \mathbf{R}^{n\times n}$, with eigenvalue decomposition

$$A = \sum_{i=1}^{n} \lambda_i q_i q_i^T,$$

where $\{q_1, \ldots, q_n\}$ are orthonormal, and $\lambda_1 \geq \cdots \geq \lambda_n$. Consider the problem of finding a nearest positive semidefinite matrix, *i.e.*, a matrix $\hat{A} = \hat{A}^T \succeq 0$ that minimizes $\|A - \hat{A}\|$. A solution to this problem is

$$\hat{A} = A = \sum_{i=1}^{n} \max\{\lambda_i, 0\} q_i q_i^T.$$

Thus, to get a nearest positive semidefinite matrix, you simply remove the terms in the eigenvector expansion that correspond to negative eigenvalues. The matrix $\hat{A}$ is sometimes called the *positive semidefinite part* of A.

As you might guess, the matrix $\hat{A}$ is also the nearest positive semidefinite matrix to A, as measured in the Frobenius norm.

3 Extremal gain problems

Suppose $A \in \mathbf{R}^{m \times n}$ has SVD

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T.$$

You already know that $v = v_1$ maximizes $\|Ax\|$ over all x with norm one. In other words, v_1 defines a direction of maximum gain for A. We can also find a direction of minimum gain. If $r < n$, then any unit vector x in $\mathcal{N}(A)$ minimizes $\|Ax\|$. If $r = n$, then the vector v_n minimizes $\|Ax\|$ among all vectors of norm one.

These results can be extended to finding *subspaces* on which A has large or small gain. Let $\mathcal{V}$ be a subspace of $\mathbf{R}^n$. We define the *minimum gain* of $A \in \mathbf{R}^{m \times n}$ on $\mathcal{V}$ as $\min\{\|Ax\| \mid x \in \mathcal{V},\ \|x\| = 1\}$. We can then pose the question: find a subspace of dimension p, on which A has the largest possible minimum gain. The solution is what you'd guess, provided $p \leq r$:

$$\mathcal{V} = \mathrm{span}\{v_1, \ldots, v_p\},$$

the span of the right singular vectors associated with the p largest singular values. The minimum gain of A on this subspace is σ_p.

If $p > r$, then any subspace of dimension p intersects the nullspace of A, and therefore has minimum gain zero. So when $p > r$ you can take $\mathcal{V}$ as any subspace of dimension p; they all have the same minimum gain, namely, zero.

We can also find a subspace $\mathcal{V}$ of dimension p that has the smallest *maximum gain* of A, defined as $\max\{\|Ax\| \mid x \in \mathcal{V},\ \|x\| = 1\}$. Assuming $r = n$ (*i.e.*, A has nullspace $\{0\}$), one such subspace is

$$\mathcal{V} = \mathrm{span}\{v_{r-p+1}, \ldots, v_r\},$$

the span of the right singular vectors associated with the p smallest singular values.

We can put state these results in a more concrete form using matrices. To define a subspace of dimension p we use an orthonormal basis, $\mathcal{V} = \mathrm{span}\{q_1, \ldots, q_p\}$. Defining $Q = [q_1 \ \cdots \ q_p]$, we have $Q^T Q = I_p$, where I_p is the $p \times p$ identity matrix. We can express the minimum gain of A on $\mathcal{V}$ as

$$\sigma_{\min}(AQ).$$

The problem of finding a subspace of dimension p that maximizes the minimum gain of A can be stated as

$$\begin{array}{ll} \text{maximize} & \sigma_{\min}(AQ) \\ \text{subject to} & Q^TQ = I_p. \end{array}$$

One solution to this problem is $Q = [v_1 \ \cdots \ v_p]$.

4 Extremal trace problems

Let $A \in \mathbf{R}^{n\times n}$ be symmetric, with eigenvalue decomposition $A = \sum_{i=1}^n \lambda_i q_i q_i^T$, with $\lambda_1 \geq \cdots \geq \lambda_n$, and $\{q_1, \ldots, q_n\}$ orthonormal. You know that a solution of the problem

$$\begin{array}{ll} \text{minimize} & x^TAx \\ \text{subject to} & x^Tx = 1, \end{array}$$

where the variable is $x \in \mathbf{R}^n$, is $x = q_n$. The related maximization problem is

$$\begin{array}{ll} \text{maximize} & x^TAx \\ \text{subject to} & x^Tx = 1, \end{array}$$

with variable $x \in \mathbf{R}^n$. A solution to this problem is $x = q_1$.

Now consider the following generalization of the first problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(X^TAX) \\ \text{subject to} & X^TX = I_k, \end{array}$$

where the variable is $X \in \mathbf{R}^{n\times k}$, and I_k denotes the $k \times k$ identity matrix, and we assume $k \leq n$. (The constraint means that the columns of X are orthonormal.) A solution of this problem is $X = [q_{n-k+1} \cdots q_n]$. Note that when $k = 1$, this reduces to the first problem above.

The related maximization problem is

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}(X^TAX) \\ \text{subject to} & X^TX = I_k, \end{array}$$

with variable $X \in \mathbf{R}^{n\times k}$. A solution of this problem is $X = [q_1 \cdots q_k]$.