

CS 6210: Matrix Computations

Matrix nearness problems

David Bindel

2025-10-15

Matrix nearness problems

A matrix nearness problem has the form

$$\text{minimize } \|X - A\| \text{ s.t. } X \in \Omega$$

or, equivalently,

$$\text{minimize } \|E\| \text{ s.t. } A + E \in \Omega$$

where Ω is a set in matrix space (real or complex) and A is a target matrix. The most frequent choice of norms are the Frobenius norm and the operator 2-norm (aka the spectral norm). Depending on the context, one may be interested in simple bounds on the minimum value, an explicit formula or characterization for the minimum value, characterization of any X (or E) for which the minimum value is obtained, or an algorithm for computing or estimating either the minimum value $\|E\|$ or an explicit minimizer E .

Our treatment of matrix nearness problems is largely drawn from the excellent paper “Matrix nearness problems and applications” by Nick Higham, appearing in *Applications of Matrix Theory* (Oxford University Press, 1989).

Preliminaries

In most cases, the easiest norm to work with for matrix nearness problems is the Frobenius norm, for a few reasons:

- The squared Frobenius norm is an inner product norm with respect to the Frobenius inner product, and is everywhere differentiable (in the real case), with

$$\delta [\|A\|_F^2] = 2\langle \delta A, A \rangle_F = 2 \operatorname{tr}(A^T \delta A).$$

- The Frobenius norm is *strictly* convex. All norms are convex by homogeneity together with the triangle inequality; that is, for $0 \leq \alpha \leq 1$ we have

$$\|\alpha x + (1 - \alpha)y\| \leq \alpha\|x\| + (1 - \alpha)\|y\|.$$

But for the Frobenius norm (and the vector 2-norm), we have strict inequality when $x \neq y$ and $0 < \alpha < 1$. Strict convexity allows us to get uniqueness results for the minimizer in the Frobenius norm in some cases where we do not have uniqueness in other norms.

- The Frobenius norm is *unitarily invariant*, i.e.

$$\|PAQ\|_F = \|A\|_F$$

whenever P, Q are unitary matrices. This means in particular that we can use the SVD and related decompositions to simplify Frobenius-norm nearness problems, since if $A = U\Sigma V^*$ is a singular value decomposition for A , then $\|A\|_F = \|\Sigma\|_F$.

One sometimes sees useful nearness results with respect to general unitarily invariant norms. The most common such norms are the *Ky-Fan* norms. The Ky-Fan p norms have the form

$$\|A\| = \|\sigma\|_p$$

where σ is the vector of singular values of A ; the Frobenius norm and the spectral norm are the Ky-Fan 2-norm and the Ky-Fan ∞ -norm, respectively. The Ky-Fan 1-norm (also called the nuclear norm) is also used in some applications. However, the spectral norm and the nuclear norm lack the differentiability and strict convexity of the Frobenius norm.

Symmetry

A warm-up case is the question of the nearest symmetric matrix. The space $\mathbb{R}^{n \times n}$ of square matrices can be written as a direct sum of the $n(n+1)/2$ -dimensional space of symmetric matrices ($H = H^T$) and the $n(n-1)/2$ -dimensional space of skew matrix ($K = -K^T$). The two spaces are orthogonal to each other in the Frobenius inner product; and for any matrix $A \in \mathbb{R}^{n \times n}$, there is a unique decomposition into a symmetric and a skew symmetric part:

$$A = A_H + A_K, \quad A_H = A_H^T, \quad A_K = -A_K^T$$

where $A_H = (A + A^T)/2$ and $A_K = (A - A^T)/2$. The best symmetric approximation to A in the Frobenius norm is therefore A_H , since the residual A_K is normal to the space of symmetric matrices. And by the Pythagorean theorem, $\|A\|_F^2 = \|A_H\|_F^2 + \|A_K\|_F^2$, so $\|A_K\|_F^2 = \|A - A_H\|_F^2 = \|A\|_F^2 - \|A_H\|_F^2$ is the distance from A to the closest symmetric matrix.

What if we are interested in other norms? The characterization of the distance to symmetry is straightforward in any unitarily invariant norm: it is always $\|A - A_H\| = \|A_K\|$. To prove this,

Fan and Hoffman used the fact that unitary invariance implies that $\|A\| = \|A^T\|$, and so for any symmetric Y

$$\begin{aligned}\|A_K\| &= \frac{1}{2}\|(A - Y) + (Y^T - A^T)\| \\ &\leq \frac{1}{2}\|A - Y\| + \frac{1}{2}\|Y^T - A^T\| \\ &\leq \|A - Y\|.\end{aligned}$$

The minimum distance is achieved at $X = A_H$, but it generally may be achieved by other points, too – the uniqueness that we see in the Frobenius norm doesn't generalize. For example, consider

$$A = \begin{bmatrix} 0 & -1 & & \\ 1 & 0 & & \\ & & 0.1 & \\ & & & 0.1 \end{bmatrix}$$

The symmetric part of this matrix is $A_H = \text{diag}(0, 0, 0.1, 0.1)$, but in the spectral norm it is the same distance from A as the all zero matrix, for example: $\|A_K\|_2 = \|A\|_2 = 1$.

Distance to rank deficiency

Suppose $A \in \mathbb{R}^{n \times n}$, and consider the problem of finding the smallest E such that a given $x \neq 0$ is a null vector of $A + E$. Take any operator norm associated with some vector norm, and let z^T be a dual vector to x with respect to the vector norm (i.e. $\|z^T\| = 1$ in the appropriate dual norm and $z^T x = \|x\|$). The smallest possible $\|E\|$ in the operator norm is $\|Ax\|/\|x\|$, and this is attained at $E = -Axz^T$. Now, if we minimize $\|Ax\|/\|x\|$ over all nonzero x , the minimum possible value is $\|A^{-1}\|^{-1}$, which gives us that

$$\min \left\{ \frac{\|E\|}{\|A\|} : A + E \text{ is singular} \right\} = \kappa(A)^{-1}.$$

That is, the inverse condition number can be seen as the relative distance to singularity of the matrix A , giving us a nice geometric interpretation of the condition number (and this geometric interpretation extends to many other settings).

Low rank and Eckart-Young-Mirsky

Closely related to the distance of a square matrix to the nearest singular matrix is the problem of distance to rank deficiency for a possibly rectangular $A \in \mathbb{R}^{m \times n}$. Then the minimum distance to a rank k matrix is achieved by the truncated SVD:

$$A_k = U_k \Sigma_k V_k^T$$

where U_k and V_k consist of the first k columns of the singular vector matrices U and V , and Σ_k is the diagonal matrix of the k largest singular values. In the Frobenius norm, this was proved by Eckart and Young, and it was later shown true in any unitarily invariant norm – hence it is called the Eckart-Young-Mirsky theorem. We will discuss the Frobenius norm case.

Suppose $\|A - B\|_F^2$ is minimal, where $B = XDY^T$, $X, Y \in \mathbb{R}^{n \times k}$ have orthonormal columns and $D \in \mathbb{R}^{k \times k}$ is diagonal with non-negative entries. Note that we can allow X and Y to deviate from having orthonormal columns, but there will always exist *some* representation of the stated form (by the SVD). Expanding the quadratic and playing with the cyclic property of traces gives

$$\begin{aligned}\phi(X, D, Y) &= \|A - XDY^T\|_F^2 \\ &= \|A\|_F^2 - 2 \operatorname{tr}(A^T X D Y^T) + \|X D Y^T\|_F^2 \\ &= \|A\|_F^2 - 2 \operatorname{tr}(Y^T A X D) + \operatorname{tr}(Y^T Y D X^T X D) \\ &= \|A\|_F^2 - 2 \operatorname{tr}(X^T A^T Y D) + \operatorname{tr}(X^T X D Y^T Y D)\end{aligned}$$

Differentiating with respect to X , Y , and D gives

$$\delta\phi = 2\langle D - Y^T A X, \delta D \rangle_F + 2\langle (Y D - A X) D, \delta Y \rangle_F + 2\langle (X D - A^T Y) D, \delta X \rangle_F$$

Setting the gradient to zero, we have the stationary conditions

$$\begin{aligned}D &= \operatorname{diag}(Y^T A X) \\ (Y D - A X) D &= 0 \\ (X D - A^T Y) D &= 0\end{aligned}$$

If $d_j > 0$, then the latter two equations give

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} y_j \\ x_j \end{bmatrix} = \begin{bmatrix} y_j \\ x_j \end{bmatrix} d_j,$$

i.e. the columns of A solve an eigenvalue problem. In fact, as we will see after the fall break, the solutions to this eigenvalue problem with positive eigenvalues are exactly (up to choice of normalization)

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} v_j \\ u_j \end{bmatrix} = \begin{bmatrix} v_j \\ u_j \end{bmatrix} \sigma_j.$$

Therefore, the columns of X and Y must *either* satisfy $y_j^T A x_j = d_j = 0$ (in which case they really contribute nothing to B) or they must correspond to the singular vectors. Given this, we have that at a stationary point, $U^T(A - B)V$ is a diagonal matrix of singular values with k of them “zeroed out”; the best choice to zero out in order to minimize $\|A - B\|_F = \|U^T(A - B)V\|_F$ is obviously the k largest.

We will discuss Eckart-Young-Mirsky in more detail after the break, when we talk about eigenvalue problems and the singular value decomposition.

Nearest symmetric positive semidefinite

Now consider the problem of finding the nearest symmetric positive definite X to a given A . Taking the symmetric/skew symmetric decomposition of $A = A_H + A_K$, we have

$$\|A - X\|_F^2 = \|A_H - X\|_F^2 + \|A_K\|_F^2;$$

that is, we can just focus on the X that is nearest to the symmetric matrix A_H . Take the symmetric eigenvalue decomposition $A_H = Q\Lambda Q^T$, and let $\tilde{X} = Q^T X Q$; then we seek to minimize $\|A - \tilde{X}\|_F^2$ subject to the constraint that $\tilde{X}$ is positive semidefinite. A positive semidefinite matrix must have non-negative diagonal entries, so the best choice we can make is to have $\tilde{X}$ be a diagonal matrix with entries $\max(\lambda_i, 0)$.

Orthogonal nearness

We begin this section with a matrix decomposition closely related to the SVD: the so-called *polar decomposition*. Suppose $A \in \mathbb{R}^{m \times n}$, and consider the economy SVD $A = U\Sigma V^T$. We can rewrite this as

$$A = (UV^T)(V\Sigma V^T) = QH$$

where $Q = UV^T$ has orthonormal columns and $H = V\Sigma V^T$ is symmetric and positive (semi)definite. This gives us a generalization of writing a vector as a unit vector times a non-negative length.

Now suppose that $A = U\Sigma V^T$ and we want to find the closest orthogonal matrix to A in the Frobenius norm. That is, we seek W with orthonormal columns so as to minimize

$$\|A - W\|_F^2 = \|A\|_F^2 - 2\operatorname{tr}(W^T A) + \|W\|_F^2$$

Note that $\|W\|_F^2 = \sqrt{n}$ by the assumption that W has orthonormal columns, so minimizing $\|A - W\|_F$ is equivalent to maximizing (using the cyclic property of traces)

$$\operatorname{tr}(W^T A) = \operatorname{tr}(\Sigma V^T W^T U) = \langle (WV)^T U, \Sigma \rangle_F.$$

This is the same as the sum of the dot products of columns of WV^T and columns of U , weighted by Σ . These column dot products of unit vectors have maximal value of 1, taken on when the two arguments are equal; that is, we require $WV = U$ or $W = UV^T = Q$.

A closely related problem is the *orthogonal Procrustes* problem: for $A, B \in \mathbb{R}^{m \times n}$, find the minimum of $\|A - BQ\|_F$ where $Q \in \mathbb{R}^{n \times n}$ is orthogonal. As before, we note that

$$\|A - BQ\|_F^2 = \|A\|_F^2 - 2\operatorname{tr}(A^T BQ) + \|BQ\|_F^2$$

and by orthogonal invariance, $\|BQ\|_F^2 = \|B\|_F^2$ is independent of Q . Therefore, minimizing $\|A - BQ\|_F^2$ is equivalent to maximizing

$$\operatorname{tr}(A^T BQ) = \langle B^T A, Q \rangle_F.$$

Therefore, we need Q to be the polar factor of $B^T A$ with orthonormal columns.