ARTICLE

Siddharth Joshi and Stephen Boyd
Subspaces that Minimize the Condition Number of a Matrix
***Rejecta Mathematica*, Vol. 1, No. 1, pp. 4-9, July 2009**
2009 Rejecta Publications
math.rejecta.org

An open letter concerning

Subspaces that Minimize the Condition Number of a Matrix

Siddharth Joshi Stephen Boyd

This article poses and answers the following question: How do you choose a subspace of given dimension that minimizes the condition number of a given matrix on that subspace? Part of the answer is a bit surprising (at least to us): When the subspace dimension is no more than half the size of the matrix, a subspace can be found on which the matrix has condition number one.

We think our paper makes it clear that we consider our result simple, but interesting and not obvious. We certainly make no claims as to its depth, or its potential applications. It is not in the literature, and does not follow in any direct or simple way from existing results. In other words, it is, as far as we know, new.

The manuscript was rejected by two journals. The first rejection was based on the reviewers and editor noting that someone had written a paper that seemed to cover similar material. But a cursory reading of that paper, and ours, show that while the other paper shared a few key words with ours, the results were in no way related. On the positive side, one reviewer suggested a simplification of our proof, which we gladly used in our revision, which was also rejected.

We then submitted the article to another journal. In this case, the editor apparently did not even understand the result, which is stated very clearly, in completely standard, and elementary, mathematical language. Moreover, he insisted that we describe an application, so we added a simple application involving an ellipsoid intersected with a subspace. It was rejected.

Subspaces that Minimize the Condition Number of a Matrix

Siddharth Joshi Stephen Boyd*

Abstract

We define the condition number of a nonsingular matrix on a subspace, and consider the problem of finding a subspace of given dimension that minimizes the condition number of a given matrix. We give a general solution to this problem, and show in particular that when the given dimension is less than half the dimension of the matrix, a subspace can be found on which the condition number of the matrix is one.

1 The problem

Suppose $A \in \mathbf{R}^{n\times n}$ and $\mathcal{V} \subseteq \mathbf{R}^n$ is a subspace with $\dim \mathcal{V} = k \geq 1$. We define the *maximum gain* (*minimum gain*) of A on $\mathcal{V}$, as

$$G_{\max} = \sup_{x\in\mathcal{V},\ x\neq 0} \frac{\|Ax\|}{\|x\|}, \qquad G_{\min} = \inf_{x\in\mathcal{V},\ x\neq 0} \frac{\|Ax\|}{\|x\|},$$

respectively, where $\|\ \|$ denotes the Euclidean norm. When A is nonsingular, we define its *condition number on the subspace* $\mathcal{V}$ as

$$\kappa_{\mathcal{V}}(A) = G_{\max}/G_{\min}.$$

The condition number of A on any one-dimensional subspace is 1, and its condition number on $\mathcal{V} = \mathbf{R}^n$ is the (usual) condition number of A, which we denote $\kappa(A)$. The condition number on any subspace is between 1 and $\kappa(A)$. If $\kappa_{\mathcal{V}}(A) = 1$, we say that A is isotropic on $\mathcal{V}$, since its gain $\|Ax\|/\|x\|$ is the same for any nonzero vector $x \in \mathcal{V}$.

In this note we address the following problem: Given a nonsingular matrix $A \in \mathbf{R}^{n\times n}$, and $k \in \{1, \ldots, n\}$, find a subspace $\mathcal{V} \subseteq \mathbf{R}^n$ of dimension k which minimizes $\kappa_{\mathcal{V}}(A)$. The number $\kappa_{\mathcal{V}}(A)$ is a measure of the anisotropy of the linear function induced by A, restricted to the subspace $\mathcal{V}$, so our problem is to find a subspace of dimension k on which A is maximally isotropic.

*The authors are with the department of Electrical Engineering at Stanford University. Email addresses: Siddharth Joshi: `sidj@stanford.edu`, Stephen Boyd: `boyd@stanford.edu`.

We will show that the minimum possible condition number of A, on a subspace of dimension k, is given by

$$\inf_{\mathcal{V}\,:\,\dim\mathcal{V}=k} \kappa_{\mathcal{V}}(A) = \max\left(\frac{\sigma_{n-k+1}}{\sigma_k}, 1\right) = \begin{cases} 1 & k \leq \lceil n/2 \rceil, \\ \sigma_{n-k+1}/\sigma_k & k > \lceil n/2 \rceil, \end{cases} \qquad (1)$$

where $\sigma_1 \geq \cdots \geq \sigma_n > 0$ are the singular values of A. (The infimum is over all subspaces of $\mathbf{R}^n$ of dimension k.) This means, in particular, that for $k \leq \lceil n/2 \rceil$, we can find a subspace of dimension k on which A is isotropic.

There are many classical results that identify a subspace of a given dimension that minimizes or maximizes some quantity that depends on the subspace and matrix. For example, the Courant-Fischer theorem tells us that the minimum value of $G_{\max}$, over all subspaces of dimension k, is σ_{n-k+1}, and the maximum value of $G_{\min}$, over all subspaces of dimension k, is σ_k. For these and similar results, see, e.g., [3, §4.2] or [1]. Also, the idea of condition number of a matrix restricted to a particular subspace can be seen in [2].

We can give a geometric application (or interpretation) of our problem. We are given an ellipsoid $\mathcal{E} = \{z \mid \|Az\| \leq 1\}$ in $\mathbf{R}^n$, where $A \in \mathbf{R}^{n\times n}$ is nonsingular. Our goal is to find a k dimensional subspace $\mathcal{V}$ so that the ellipsoid $\mathcal{V} \cap \mathcal{E}$ is as spherical as possible, *i.e.*, has minimum eccentricity. (The eccentricity of $\mathcal{V} \cap \mathcal{E}$ is defined as the ratio of its maximum semi-axis length to its minimum semi-axis length, which is exactly $\kappa_{\mathcal{V}}(A)$.) The solution is to choose $\mathcal{V}$ that minimizes the condition number of A on $\mathcal{V}$. Our result (1) can be interpreted in this geometric setting. For example, if $k < \lceil n/2 \rceil$, we can always find a subspace of dimension k for which $\mathcal{V} \cap \mathcal{E}$ is perfectly spherical, *i.e.*, a ball. As a very simple special case, we see that for any ellipsoid in $\mathbf{R}^3$, there is a plane that intersects it in a ball. Our general result (1) can be considered a generalization of this simple fact.

2 The solution

Suppose Q and Z are $n \times n$ orthogonal matrices, *i.e.*, $Q^TQ = Z^TZ = I$. Then we have

$$\kappa_{\mathcal{V}}(QAZ) = \kappa_{\mathcal{W}}(A),$$

where $\mathcal{W} = Z\mathcal{V} = \{Zv \mid v \in \mathcal{V}\}$. It follows that

$$\inf_{\mathcal{V}\,:\,\dim\mathcal{V}=k} \kappa_{\mathcal{V}}(A) = \inf_{\mathcal{V}\,:\,\dim\mathcal{V}=k} \kappa_{\mathcal{V}}(QAZ),$$

since the first orthogonal matrix Q has no effect, and the second orthogonal matrix Z simply changes the parametrization of subspaces of dimension k.

Now let $A = U\Sigma V^T$ be a singular value decomposition of A, *i.e.*, U and V are orthogonal, and $\Sigma = \mathbf{diag}(\sigma_1, \ldots, \sigma_n)$. Our observation above, with $Q = U^T$, $Z = V$, shows that

$$\inf_{\mathcal{V}\,:\,\dim\mathcal{V}=k} \kappa_{\mathcal{V}}(A) = \inf_{\mathcal{V}\,:\,\dim\mathcal{V}=k} \kappa_{\mathcal{V}}(\Sigma).$$

So we can just as well solve the problem for the diagonal matrix Σ. (To reconstruct a subspace of dimension k on which A has least condition number, we find a subspace of dimension k for which Σ has least condition number, and multiply it by V.)

Now our problem is to find a subspace $\mathcal{V}$ of dimension k which minimizes $\kappa_{\mathcal{V}}(\Sigma)$. We will show that

$$\inf_{\mathcal{V}\,:\,\dim\mathcal{V}=k} \kappa_{\mathcal{V}}(\Sigma) = \max\left(\frac{\sigma_{n-k+1}}{\sigma_k}, 1\right) = \begin{cases} 1 & k \leq \lceil n/2 \rceil, \\ \sigma_{n-k+1}/\sigma_k & k > \lceil n/2 \rceil, \end{cases} \tag{2}$$

Let $\{e_1, \ldots, e_n\}$ be the standard basis for $\mathbf{R}^n$, *i.e.*, for $i = 1, \ldots, n$, $e_{ij} = 0$ if $i \neq j$ and $e_{ij} = 1$ otherwise.

We first give a simple result. Suppose $i < j$, and let σ satisfy $\sigma_i \geq \sigma \geq \sigma_j$. Then there is a unit vector $z \in \mathbf{span}\{e_i, e_j\}$ for which $\|\Sigma z\| = \sigma$. This can be seen several ways. For example, we can rotate a unit vector z from e_i towards e_j. The norm $\|\Sigma z\|$ varies continuously from σ_i to σ_j, and therefore has the value σ at some rotation angle. We can easily construct such a z. If $\sigma_i = \sigma_j$, we can take $z = e_i$ or $z = e_j$. If $\sigma_i > \sigma_j$, we can take

$$z = \frac{(\sigma^2 - \sigma_j^2)^{1/2} e_i + (\sigma_i^2 - \sigma^2)^{1/2} e_j}{(\sigma_i^2 - \sigma_j^2)^{1/2}}.$$

It is easily verified that $\|z\| = 1$ and $\|\Sigma z\| = \sigma$.

2.1 Case 1: $k \leq \lceil n/2 \rceil$

To establish (2), we will construct a subspace $\mathcal{V}^*$ of dimension k, with $\kappa_{\mathcal{V}^*}(\Sigma) = 1$. We will construct an orthonormal basis $\{z_0, z_1, \ldots, z_{k-1}\}$ for $\mathcal{V}^*$. We start with $z_0 = e_{\lceil n/2 \rceil}$. Note that $\|\Sigma z_0\| = \sigma_{\lceil n/2 \rceil}$.

Next, we choose a unit vector $z_1 \in \mathbf{span}\{e_{\lceil n/2 \rceil - 1}, e_{\lceil (n+1)/2 \rceil + 1}\}$ that satisfies $\|\Sigma z_1\| = \sigma_{\lceil n/2 \rceil}$. We can do this using our simple result above, noting that

$$\sigma_{\lceil n/2 \rceil - 1} \geq \sigma_{\lceil n/2 \rceil} \geq \sigma_{\lceil (n+1)/2 \rceil + 1}.$$

We note that $z_1 \perp z_0$ and $\Sigma z_1 \perp \Sigma z_0$.

We continue the construction, taking z_2 as any unit vector

$$z_2 \in \mathbf{span}\{e_{\lceil n/2 \rceil - 2}, e_{\lceil (n+1)/2 \rceil + 2}\}$$

that satisfies $\|\Sigma z_2\| = \sigma_{\lceil n/2 \rceil}$. This continues, until we have unit vectors $z_0, \ldots, z_{k-1}$. These vectors are mutually orthogonal, since each one is in the span of two standard basis vectors, and these pairs of standard basis vectors are disjoint. Since Σ is a diagonal matrix, the vectors $\Sigma z_0, \ldots, \Sigma z_{k-1}$ are mutually orthogonal.

We now show that $\kappa_{\mathcal{V}^*}(\Sigma) = 1$. For any nonzero vector $b \in \mathcal{V}^*$, the gain of Σ in the direction of b, $\|\Sigma b\|/\|b\| = \sigma_{\lceil n/2 \rceil}$, because the gain of Σ in the direction of any unit vector in the orthonormal basis $\{z_0, \ldots, z_{k-1}\}$ of $\mathcal{V}^*$ is $\sigma_{\lceil n/2 \rceil}$. Thus $G_{\max} = G_{\min} = \sigma_{\lceil n/2 \rceil}$, and therefore $\kappa_{\mathcal{V}^*}(\Sigma) = 1$.

2.2 Case II: $k > \lceil n/2 \rceil$

To establish (2), we first construct a subspace $\mathcal{V}^*$ of dimension k, with $\kappa_{\mathcal{V}^*}(\Sigma) = \sigma_{n-k+1}/\sigma_k$, and then show that for any subspace $\mathcal{V}$ of dimension k, $\kappa_{\mathcal{V}}(\Sigma) \geq \kappa_{\mathcal{V}^*}(\Sigma)$.

We will construct an orthonormal basis for $\mathcal{V}^*$. We start with the $2k - n$ vectors $\{e_{n-k+1}, e_{n-k}, \ldots, e_{k-1}, e_k\}$. We will choose $n - k$ unit vectors, $z_1, \ldots, z_{n-k}$, such that

$$\{z_1, \ldots, z_{n-k}, e_{n-k+1}, \ldots, e_{k-1}, e_k\}$$

forms an orthonormal basis for $\mathcal{V}^*$. The $n-k$ unit vectors $z_1, \ldots, z_{n-k}$ will be chosen in $\mathbf{span}\{e_1, \ldots, e_{n-k}, e_{k+1}, \ldots, e_n\}$, and will therefore be orthogonal to $\{e_{n-k+1}, \ldots, e_k\}$.

Choose a unit vector $z_1 \in \mathbf{span}\{e_1, e_n\}$, satisfying $\|\Sigma z_1\| = \sigma_k$. We can do this using the simple result given earlier, since $\sigma_1 \geq \sigma_k \geq \sigma_n$. We note that $z_1 \perp e_j$, and $\Sigma z_1 \perp \Sigma e_j$, $j = n-k+1, \ldots, k$.

We continue the construction, choosing a unit vector $z_2 \in \mathbf{span}\{e_2, e_{n-1}\}$, satisfying $\|\Sigma z_2\| = \sigma_k$. This continues, until we have chosen a unit vector z_{n-k} in $\mathbf{span}\{e_{n-k}, e_{k+1}\}$, satisfying $\|\Sigma z_{n-k}\| = \sigma_k$.

The vectors $z_1, \ldots, z_{n-k}$ are mutually orthogonal, since each one is in the span of two standard basis vectors, and these pairs of standard basis vectors are disjoint. Also $z_i \perp e_j$ for $i = 1, \ldots, n-k$ and $j = n-k+1, \ldots, k$, since each vector z_i is in the span of two standard basis vectors which are not in the set $\{e_{n-k+1}, \ldots, e_k\}$. Thus $\{z_1, \ldots, z_{n-k}, e_{n-k+1}, e_{n-k}, \ldots, e_{k-1}, e_k\}$ forms an orthonormal basis for $\mathcal{V}^*$. Similarly, since Σ is a diagonal matrix, the vectors $\Sigma z_1, \ldots, \Sigma z_{n-k}, \Sigma e_{n-k+1}, \ldots, \Sigma e_k$ are mutually orthogonal.

We now show $\kappa_{\mathcal{V}^*}(\Sigma) = \sigma_{n-k+1}/\sigma_k$. Let b any nonzero vector in $\mathcal{V}^*$, say,

$$b = \beta_1 z_1 + \cdots + \beta_{n-k} z_{n-k} + \beta_{n-k+1} e_{n-k+1} + \cdots + \beta_k e_k.$$

The gain of Σ in the direction b is

$$\begin{aligned} \frac{\|\Sigma b\|}{\|b\|} &= \left(\frac{\sum_{i=1}^{n-k} \beta_i^2 \|\Sigma z_i\|^2 + \sum_{j=n-k+1}^{k} \beta_j^2 \|\Sigma e_j\|^2}{\sum_{i=1}^{n-k} \beta_i^2 \|z_i\|^2 + \sum_{j=n-k+1}^{k} \beta_j^2 \|e_j\|^2} \right)^{1/2} \\ &= \left(\frac{\sum_{i=1}^{n-k} \beta_i^2 \sigma_k^2 + \sum_{j=n-k+1}^{k} \beta_j^2 \sigma_j^2}{\sum_{i=1}^{n} \beta_i^2} \right)^{1/2}, \end{aligned}$$

and therefore $\sigma_{n-k+1} \geq \|\Sigma b\|/\|b\| \geq \sigma_k$. For $b = e_{n-k+1}$, $\|\Sigma b\|/\|b\| = \sigma_{n-k+1}$, so $G_{\max} = \sigma_{n-k+1}$; for $b = e_k$, we have $\|\Sigma b\|/\|b\| = \sigma_k$, so $G_{\min} = \sigma_k$. It follows that $\kappa_{\mathcal{V}^*}(\Sigma) = \sigma_{n-k+1}/\sigma_k$.

Now we will show that for any subspace $\mathcal{V}$ of dimension k, $\kappa_{\mathcal{V}}(\Sigma) \geq \sigma_{n-k+1}/\sigma_k$. By the Courant-Fischer theorem, for any subspace $\mathcal{V}$ of dimension k, $G_{\max} \geq \sigma_{n-k+1}$ and $G_{\min} \leq \sigma_k$. It follows that $\kappa_{\mathcal{V}}(\Sigma) = G_{\max}/G_{\min} \geq \sigma_{n-k+1}/\sigma_k$. This establishes (2), and therefore (1).

References

[1] D. S. Bernstein. *Matrix Mathematics: Theory, facts, and formulas, with application to linear systems theory.* Princeton University Press, 2005.

[2] T. F. Chan and D. E. Foulser. Effectively well-conditioned linear systems. *SIAM Journal on Scientific and Statistical Computing*, 9(6):963–968, November 1988.

[3] R. A. Horn and C. A. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.