RESEARCH
COUNCILS UK

nag

HECToR

CSC / NAG Autumn School on

Core Algorithms in High-Performance Scientific Computing

Maths VI

Dwight Barkley

Miscellaneous Topics

Maths VI: Miscellaneous Topics

In this lecture we consider least squares problems and the closely related problems of "solving" linear equations $Ax = b$ in the case where A is singular or nearly singular.

6.1 Introduction

An example

Probably the most familiar example of least squares is curve fitting: Finding a line, or more generally a polynomial, approximating a set of data points. More specifically, given m data points

$$(x_1, y_1), (x_2, y_2), \cdots (x_m, y_m)$$

the problem is to find n coefficients $c_0, \cdots, c_{n-1}$ of the polynomial

$$p(x) = c_0 + c_1 x + \cdots c_{n-1} x^{n-1}$$

giving the best fit to the data.

This is expressed as a linear algebra problem by evaluating the polynomial at the data points and requiring (or attempting to require) $y_j = p(x_j)$

$$y_j = c_0 + c_0 x_j + \cdots c_{n-1} x_j^{n-1}$$

This provides a set of linear conditions on the unknown coefficients in the form

$$Ac = y$$

where A is the Vandermonde matrix

$$A = \begin{bmatrix} 1 & x_1 & \ldots & x_1^{n-1} \\ 1 & x_2 & \ldots & x_2^{n-1} \\ \vdots & & & \\ 1 & x_m & \ldots & x_m^{n-1} \end{bmatrix}$$

If $n = m$ and the points are distinct, then the matrix A square and non-singular and the linear equations has a unique solution. This solution gives the unique polynomial going exactly through the data points. However, it is much more common to seek a lower order polynomial "passing through" the data. In this case $n < m$ and often n is much less than m. The matrix A is then rectangular with more rows than columns. **The system of equations is said to by over-determined** because there are more equations than unknowns. In general it is not possible to find a solution c giving exact equality. This leads to the least squares problem of finding a c such that

$$||Ac - y||_2 \quad \text{is as small as possible}$$

Note that the 2-norm is specified explicitly. The 2-norm gives a quadratic function to minimize. When differentiated to find the minimum, this gives a linear condition and so this problem is often referred to as **linear least squares (LLS)**. We shall drop the qualifier linear, but we shall always include the 2-norm in expressions.

Formal statement of the LLS problem

Given an $m \times n$ matrix A with $m > n$ and an n-vector b, find the m-vector x such that $||b - Ax||_2$ is minimized.

Since $r = b - Ax$ is the residual, we seek x such as to minimize the residual norm, in the 2-norm. One typically is interested also in $y = Ax$.

The general setting

There are a number of closely related problems:

1. **Overdetermined:** More equations than unknowns, $m > n$

$$\begin{bmatrix} \\ \quad A \quad \\ \\ \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$

2. **Underdetermined:** Fewer equations than unknowns, $m < n$

$$\begin{bmatrix} \qquad A \qquad \end{bmatrix} \begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$$

3. **Singular square systems:** $m = n$, but A^{-1} does not exist

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$

4. **Overdetermined, rank deficient:** More equations than unknowns, $m > n$, but least square solution not unique.

5. **Square, nearly singular systems:** $m = n$ and technically A^{-1} exists, but the problem is very poorly conditioned.

6.2 What Does $Ax \to y$ Look Like?

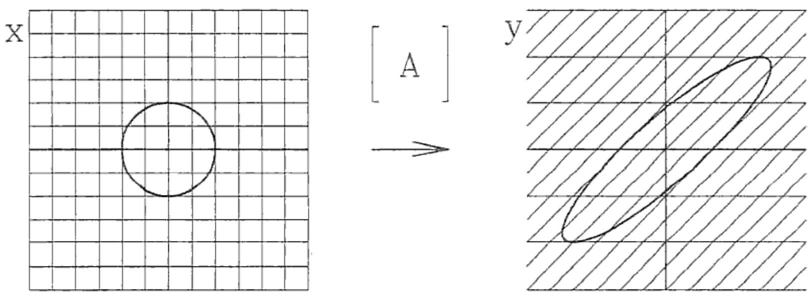


Figure 6.1: The geometry of matrix vector multiplication for a generic, non-singular square matrix. The unit sphere $||x||_2 = 1$ is mapped by A to a hyperellipse. What we will want to do is examine the corresponding picture corresponding to the various cases just listed.

6.3 An Aside on Matrix Norms

Recall the 3 most important vector norms:

- **1-norm**: $||x||_1 = \sum_{i=1}^{n} |x_i|$.
- **2-norm or Euclidean norm**: $||x||_2 = (\sum_{i=1}^{n} |x_i|^2)^{1/2} = \sqrt{x^* x}$.
- **∞-norm**: $||x||_\infty = \max_{1 \le i \le n} |x_i|$

and recall that corresponding to each vector norm, the induced matrix norm is defined as

$$||A|| = \sup_{||x||=1} ||Ax||$$

The meaning of these norms is illustrated for the case

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix}$$

The norms of A are:

$$\begin{aligned} &1 \text{ norm:} && ||A||_1 = 4 \\ &2 \text{ norm:} && ||A||_2 = 2.9208\ldots \\ &\infty \text{ norm:} && ||A||_\infty = 3 \end{aligned}$$

The 1-norm and ∞-norm of a matrix are readily computable, but not the 2-norm.

An important message, however, is that there is little qualitative difference between these cases. We will only consider the 2-norm from the remainder of these lectures.

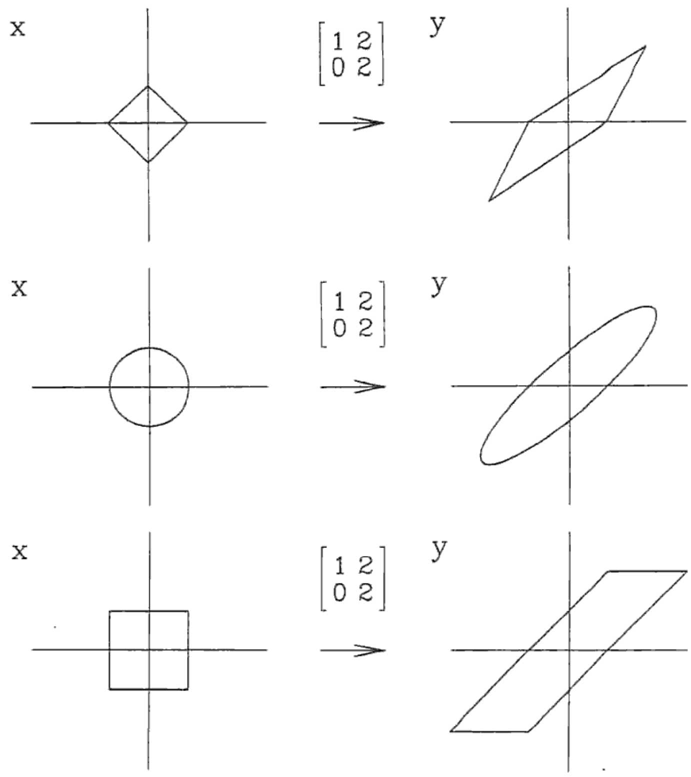


Figure 6.2: The geometry of vector and matrix norms. From top to bottom, the 1-norm, the 2-norm, and the ∞-norm. Left column shows the set of points with $||x|| = 1$ in the corresponding norm. The right shows where these points get mapped by A. The norm of the matrix A is defined to be the maximum distance from the origin of the blue points after they are mapped. Distance means as measured in the corresponding vector norm.

6.4 A Few Definitions

- **column rank** of a matrix A is the number of linearly independent columns.
- **row rank** of a matrix A is the number of linearly independent row.
- The column rank and row rank are necessarily the same and so we define the **rank of A, written $rank(A)$, to be the number of linearly independent rows or columns.**
- A rectangular matrix can have at most $rank(A) = min(m, n)$. If is has such rank it is said to have **full rank**. If A does not have full rank it is said to be **rank deficient**.
- **The range of A, written** $range(A)$ is the set of all vectors that can be expressed as Ax for some x.

- **The null space of A, written** $null(A)$ is the set of all vectors such that $Ax = 0$.

6.5 The Geometry of Our Special Problems

We will now look at the cases listed at the beginning except the last. We are interested in mapping pictures for $m \times n$ matrices. Such matrices map n-vectors to m-vectors. Hence, the x vectors have length n and the y vectors have length m.

Overdetermined More equations than unknowns, $m > n$

$$\begin{bmatrix} \\ & A & \\ \\ \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$

Underdetermined:

Fewer equations than unknowns, $m < n$

$$\begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$$

Singular square systems: $m = n$, but A^{-1} does not exist

$$\begin{bmatrix} \\ & A & \\ \\ \end{bmatrix} \begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$

Overdetermined, rank deficient: More equations than unknowns, $m > n$, but least square solution not unique.

$$\begin{bmatrix} \\ & A & \\ \\ \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$

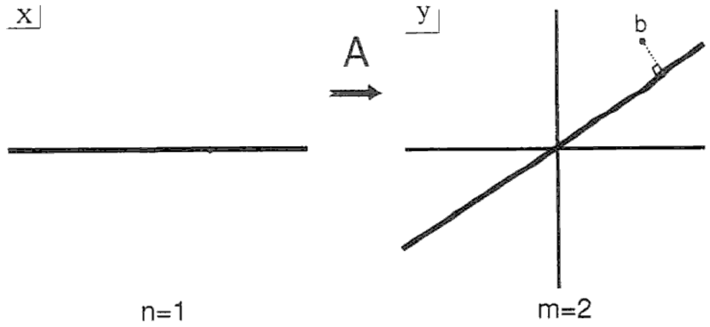


Figure 6.3: Illustration of an overdetermined system. The red line on the right indicates $range(A)$, the set of all points of the form Ax for some x. b is not in the range and so there is no x such that $Ax = b$. The least squares solution is the x (turquoise) such that Ax is closest to b, thereby minimizing the residual norm $||Ax - b||_2$. Note: the solution is given by orthogonally projecting b onto the range of A and then taking the inverse of the projected point. This is known as taking the pseudoinverse: $x = A^+$ A^+ is the **pseudoinverse** of A.

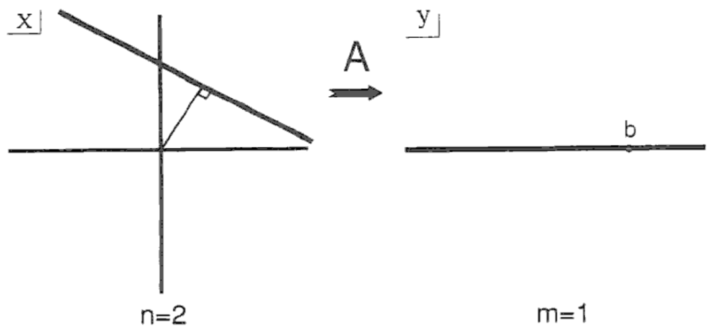


Figure 6.4: Illustration of an underdetermined system. The range of A is again shown as red on the right and is all of $\mathbb{R}^{m=1}$. In this case there is no problem finding a solution, the problem is that there are too many. It is common, and LAPACK for example will compute, the particular solution x such that $Ax = b$ and $||x||_2$ is smallest. This is called the **minimum norm solution**.

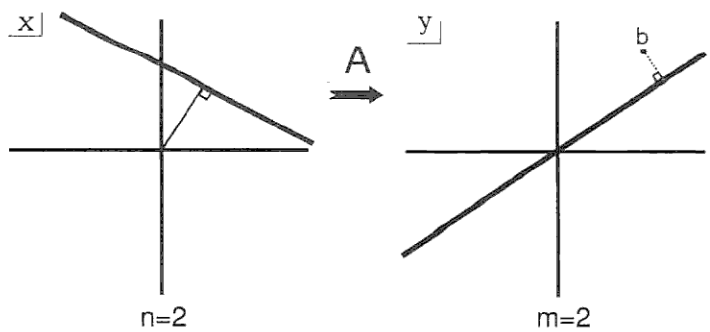


Figure 6.5: Illustration of a singular system. The range of A is again shown as red on the right. $Rank(A) = 1 < min(m, n) = 2$ so the matrix rank deficient (does not have full rank). No solutions to $Ax = b$ exists in general. One may find a least squares solution which minimizes the residual norm $||Ax - b||_2$. Here there are infinitely many x that minimize this norm. One orthogonally projects b onto the range of A and then finds the x with minimal norm such that Ax equals the projected b. This is known as the **minimal norm least squares solution**.

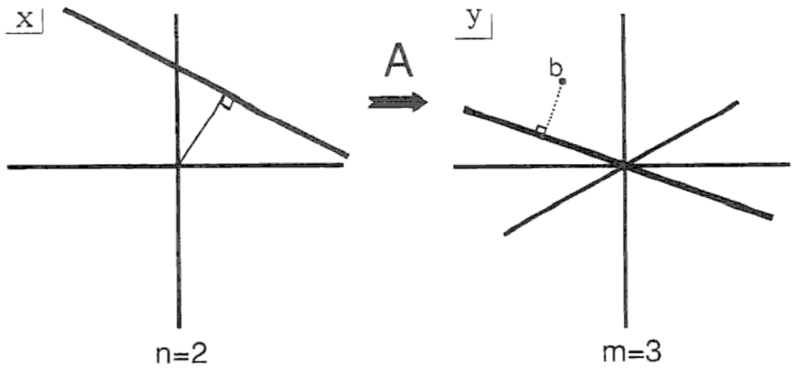


Figure 6.6: Illustration of an overdetermined system with less than full rank. The range of A is again shown as red on the right. If A had full rank the rank would be a 2 dimensional sheet. As with the singular square system, minimal norm least squares solution.

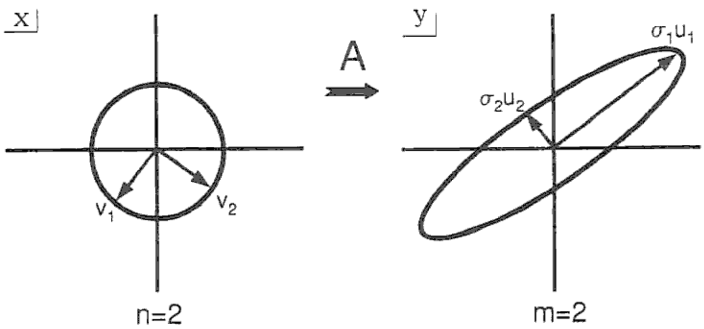


Figure 6.7: Illustration of the SVD for a square matrix. The unit sphere $||x||_2 = 1$ is mapped by A to a hyperellipse. The left singular vectors u_1 and u_2 are on the right, and right singular vectors v_1 and v_2 are on the left as shown.

6.6 The Singular Value Decomposition - SVD

Consider the case where A is $n \times n$ and non-singular. These assumptions are not essential and will be removed very quickly. Consider as before the mapping of the unit sphere by A. The image of the sphere is a hyperellipse. Singular values: The lengths of the semiaxes of the hyperellipse ordered by size.

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0 \tag{6.1}$$

Left singular vectors: u_j, $j = 1, \cdots, n$ are unit vectors along the semi axes of the hypersphere, ordered by the corresponding singular value.

Right singular vectors: v_j, $j = 1, \cdots, n$ are unit vectors such that

$$Av_j = \sigma_j u_j, \quad j = 1, \cdots, n$$

so that the right singular vectors are preimages of the left singular vectors.

Arrange the v's and u's as columns of matrices V and U. Then we have

$$AV = U\Sigma$$

where Σ is the diagonal matrix $\Sigma = diag(\sigma_1, \cdots, \sigma_n)$.

The u's and v's are each orthonormal sets of vectors: This means that the both matrices V and U are unitary (orthogonal if real)

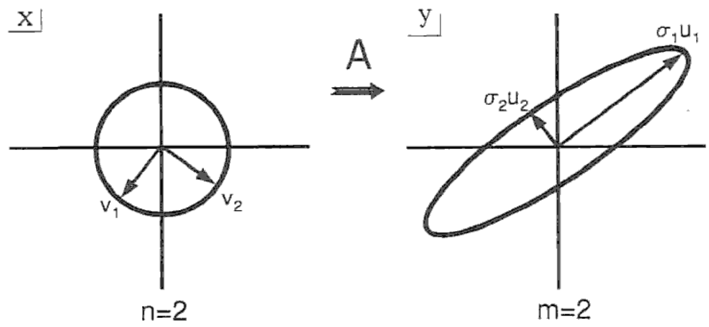
$$U^*U = I \qquad V^*V = I$$

Using this to move V^* to the right-hand-side we have:

$$A = U\Sigma V^*$$

This is the **Singular Value Decomposition SVD of matrix** A.

This decomposition is not only a useful computational tool, it provides a clear decomposition of acting with A.



- The orthogonal matrix V^* is a rotation that aligns the basis of right singular vectors with the standard basis, e.g. $V^*v_1 = e_1$.
- The diagonal matrix Σ stretches ($\sigma > 1$) or contracts ($\sigma < 1$) in the directions of the standard basis, e.g. $e_1 \to \sigma_1 e_1$.
- The orthogonal matrix U is a rotation that takes standard basis to the left singular vectors, e.g. $Ue_1 = u_1$.

Inverting is easy. $Ax = b$ gives $b = V\Sigma^{-1}U^*x$. Just the reverse of the preceding steps.

Now we may drop the assumption that A is square and non-singular. Let A be any matrix, rectangular, rank deficient it does not matter. Then there exists a singular value decomposition of A of the form:

$$A = U\Sigma V^*$$

where if A is $m \times n$ then U is $m \times m$ unitary (orthogonal if real), Σ is real diagonal with non-negative entries, and V is $n \times n$ unitary (orthogonal if real). This is known as the full SVD. There is also a reduced SVD which drops "silent" columns from the decomposition.

We can get to the essential features just from the following observations. Let $r = rank(A)$. There there will be exactly r non-zero singular values $\sigma_1, \sigma_2, \ldots, \sigma_r$. The corresponding left singular vectors $u_1, u_2, \ldots, u_r$ span $range(A)$. The corresponding right singular vectors $v_1, v_2, \ldots, v_r$ are all orthogonal to $null(A)$, with any remaining vectors $v_{r+1}, \ldots, v_n$ forming an orthogonal basis for $null(A)$. Consider the following matrices

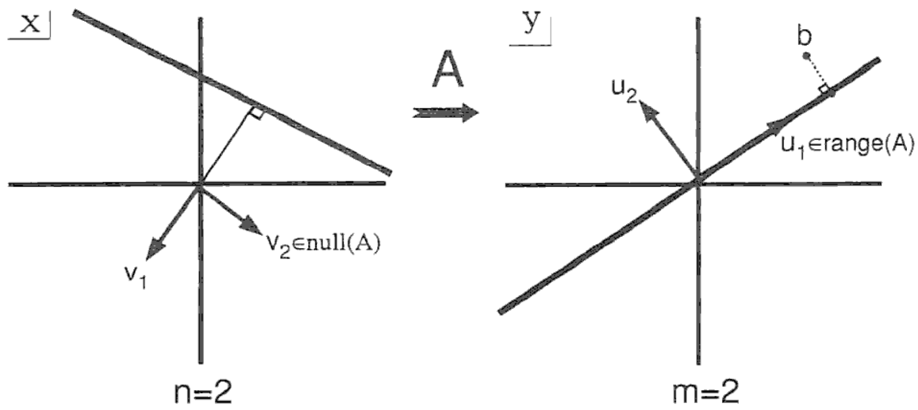
$$\tilde{U} = \begin{pmatrix} \\ u_1 & u_2 & \ldots & u_r \\ \\ \end{pmatrix} \qquad \tilde{\Sigma} = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix} \qquad \tilde{V} = \begin{pmatrix} \\ v_1 & v_2 & \ldots & v_r \\ \\ \end{pmatrix}$$

This notation is non-standard. We use it here to keep the discussion short and to the point. These matrices provide a decomposition that will give the desired solution to the 4 problems previously consider. In all cases

$$x = \tilde{V}\tilde{\Sigma}^{-1}\tilde{U}^{*}b$$

- $\tilde{U}^{*}b$: Project b onto the range of A. More accurately, find the components of b in an orthogonal basis for $range(A)$.
- $\tilde{\Sigma}^{-1}\,(\tilde{U}^{*}b)$: Multiply each component of the projected b by the inverse singular value to invert the stretching/contraction produce by A.
- $V\,(\tilde{\Sigma}^{-1}\,\tilde{U}^{*}b)$: Use the stretched/contracted components as weights for an orthogonal basis set in the complement of $null(A)$.

For A an $m \times n$ matrix with $m > n$, the "inverse matrix" "constructed" this way is known as the pseudoinverse and is denoted A^{+}.



r = rank(A) =1

In practice the LLS problem is solved either via a QR decomposition or via SVD. Read the LAPACK manual.

6.7 Nearly Singular Square Matrices

The condition number of an $n \times n$ matrix A, in the 2-norm, is the ratio of the largest singular value to the smallest singular value:

$$\kappa(A) = \frac{\sigma_1}{\sigma_n}$$

Think of this as measuring the maximum ratio of stretching to contraction by A. (Of course is could be the ratio of least contraction to most contraction if $\sigma_1 < 1$ or it could be the ratio of most stretching to least stretching if $\sigma_n > 1$).

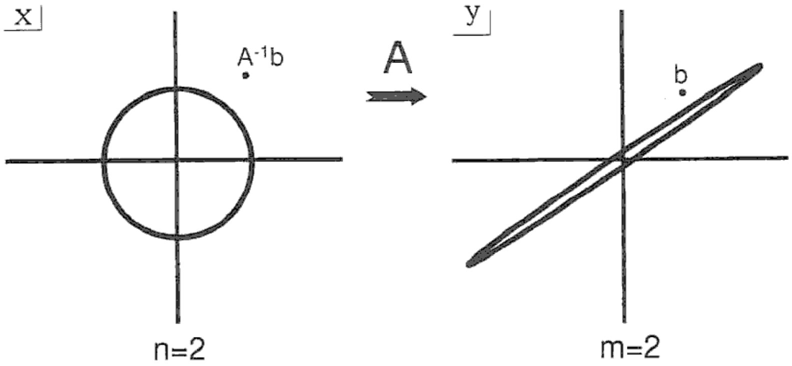


Figure 6.8: Illustration of a nearly singular matrix. The condition number is large. Even though A inverse may technically exist, solving $Ax = b$ is going to be extremely inaccurate.

One trick in the case of a poorly conditioned matrix with only a few "bad" singular values is to treat the problem as though it were rank deficient and solve a minimum norm least squares problem instead. This approach is recommended for example, in Numerical Recipes. It can certainly provide a more robust solution, although one needs to think very carefully about the meaning of such a solution and why one is trying to solve $Ax = b$ for such an A.