Linear Systems

Carlo Tomasi

Section 1 characterizes the existence and multiplicity of the solutions of a linear system in terms of the four fundamental spaces associated with the system's matrix and of the relationship of the left-hand side vector of the system to that subspace. Sections 2 and 3 then show how to use the SVD to solve a linear system in the sense of least squares. When not given in the main text, proofs are in Appendix A.

1 The Solutions of a Linear System

Let

$$A\mathbf{x} = \mathbf{b}$$

be an $m \times n$ system (m can be less than, equal to, or greater than n). Also, let

$$r = \mathrm{rank}(A)$$

be the number of linearly independent rows or columns of A. Then,

$$\begin{array}{rcl} \mathbf{b} \notin \mathrm{range}(A) & \Rightarrow & \text{no solutions} \\ \mathbf{b} \in \mathrm{range}(A) & \Rightarrow & \infty^{n-r} \text{ solutions} \end{array}$$

with the convention that $\infty^0 = 1$. Here, ∞^k is the cardinality of a k-dimensional vector space on the reals.

In the first case above, there can be no linear combination of the columns (no $\mathbf{x}$ vector) that gives $\mathbf{b}$, and the system is said to be *incompatible*. In the second, *compatible* case, three possibilities occur, depending on the relative sizes of r, m, n:

- When $r = n = m$, the system is *invertible*. This means that there is exactly one $\mathbf{x}$ that satisfies the system, since the columns of A span all of $\mathbf{R}^n$. Notice that invertibility depends only on A, not on $\mathbf{b}$.

- When $r = n$ and $m > n$, the system is *redundant*. There are more equations than unknowns, but since $\mathbf{b}$ is in the range of A there is a linear combination of the columns (a vector $\mathbf{x}$) that produces $\mathbf{b}$. In other words, the equations are compatible, and exactly one solution exists.[1]

- When $r < n$ the system is *underdetermined*. This means that the null space is nontrivial (*i.e.*, it has dimension $h > 0$), and there is a space of dimension $h = n - r$ of vectors $\mathbf{x}$ such that $A\mathbf{x} = 0$. Since $\mathbf{b}$ is assumed to be in the range of A, there are solutions $\mathbf{x}$ to $A\mathbf{x} = \mathbf{b}$, but then for any $\mathbf{y} \in \mathrm{null}(A)$ also $\mathbf{x} + \mathbf{y}$ is a solution:

  $$A\mathbf{x} = \mathbf{b}\,,\ A\mathbf{y} = 0 \ \Rightarrow\ A(\mathbf{x} + \mathbf{y}) = \mathbf{b}$$

  and this generates the $\infty^h = \infty^{n-r}$ solutions mentioned above.

[1] Notice that the technical meaning of "redundant" has a stronger meaning than "with more equations than unknowns." The case $r < n < m$ is possible, has more equations (m) than unknowns (n), admits a solution if $\mathbf{b} \in \mathrm{range}(A)$, but is called "underdetermined" because there are fewer (r) independent equations than there are unknowns (see next item). Thus, "redundant" means "with exactly one solution and with more equations than unknowns."

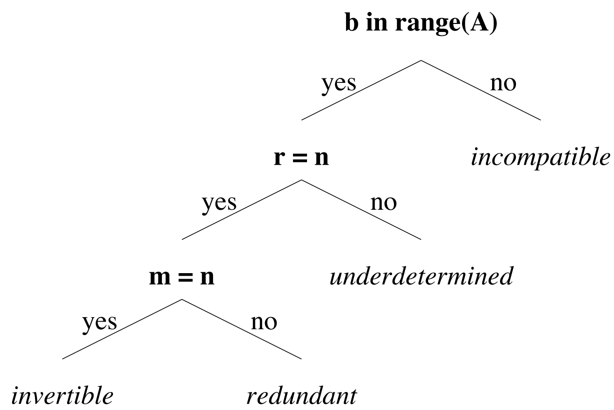

Figure 1: Types of linear systems.

Notice that if $r = n$ then n cannot possibly exceed m (or else the columns of A would form an n-dimensional subspace of an m-dimensional space with $m < n$, an impossibility), so the first two cases exhaust the possibilities for $r = n$. Also, r cannot exceed either m or n. All the cases are summarized in figure 1.

Thus, a linear system has either zero (incompatible), one (invertible or redundant), or more (underdetermined) solutions. In all cases, we can say that the set of solutions forms an *affine space*, that is, a linear space plus a vector:

$$\mathcal{A} = \hat{\mathbf{x}} + \mathcal{L} \; .$$

Recall that the sum here means that the single vector $\hat{\mathbf{x}}$ is added to every vector of the linear space $\mathcal{L}$ to produce the affine space $\mathcal{A}$. For instance, if $\mathcal{L}$ is a plane through the origin (recall that all linear spaces must contain the origin), then $\mathcal{A}$ is a plane (not necessarily through the origin) that is parallel to $\mathcal{L}$.

In the underdetermined case, the nature of $\mathcal{A}$ is obvious. The notation above also applies to the incompatible case: in this case, $\mathcal{L}$ is the empty linear space, so $\hat{\mathbf{x}} + \mathcal{L}$ is empty as well, and $\hat{\mathbf{x}}$ is undetermined.

Please do not confuse the empty linear space (a space with no elements) with the linear space that contains only the zero vector (a space with one element). The latter yields either the invertible or the redundant case.

Of course, listing all possibilities does not provide an operational method for determining the type of linear system for a given pair $A, \mathbf{b}$. The next two sections show how to do so, and how to solve the system. They also give meaning to the expression "solving the system" when no exact solution exists, which occurs most of the time in practice. Section 3, in particular, defines a concept of "solution" that is typically useful and interesting in the case $\mathbf{b} = \mathbf{0}$, when the exact solution would be trivial and uninteresting.

2 The Pseudoinverse

One of the most important applications of the SVD is the solution of linear systems in the least squares sense. A linear system of the form

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

arising from a real-life application may or may not admit a solution, that is, a vector $\mathbf{x}$ that satisfies this equation exactly. Often more measurements are available than strictly necessary, because measurements are unreliable. This leads to more equations than unknowns (the number m of rows in A is greater than the number n of columns), and equations are often mutually incompatible because they come from inexact measurements. Even when $m \leq n$ the equations can be incompatible, because of errors in the measurements that produce the entries of A. In these cases, it makes more sense to find a vector $\mathbf{x}$ that minimizes the norm

$$\|A\mathbf{x} - \mathbf{b}\|$$

of the *residual* vector

$$\mathbf{r} = A\mathbf{x} - \mathbf{b}$$

where the double bars henceforth refer to the Euclidean norm. Thus, $\mathbf{x}$ cannot exactly satisfy any of the m equations in the system, but it tries to satisfy all of them as closely as possible, as measured by the sum of the squares of the discrepancies between left- and right-hand sides of the equations.

In other circumstances, not enough measurements are available. Then, the linear system (1) is under-determined, in the sense that it has fewer independent equations than unknowns (its rank r is less than n).

Incompatibility and under-determinacy can occur together: the system admits no solution, and the least-squares solution is not unique. For instance, the system

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 + x_2 &= 3 \\ x_3 &= 2 \end{aligned}$$

has three unknowns, but rank 2, and its first two equations are incompatible: $x_1 + x_2$ cannot be equal to both 1 and 3. A least-squares solution turns out to be $\mathbf{x} = [1\ 1\ 2]^T$ with residual

$$\mathbf{r} = A\mathbf{x} - \mathbf{b} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix},$$

which has norm $\sqrt{2}$ (admittedly, this is a rather high residual, but this is the best we can do for this problem, in the least-squares sense). However, any other vector of the form

$$\mathbf{x}' = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

is as good as $\mathbf{x}$. For instance, $\mathbf{x}' = [0\ 2\ 2]$, obtained for $\alpha = 1$, yields exactly the same residual as $\mathbf{x}$ (check this).

In summary, an exact solution to the system (1) may not exist, or may not be unique. An approximate solution, in the least-squares sense, always exists, but may fail to be unique.

If there are several least-squares solutions, all equally good (or bad), then one of them turns out to be shorter than all the others, that is, its norm $\|\mathbf{x}\|$ is smallest. One can therefore redefine what it means to "solve" a linear system so that there is always exactly one solution. This minimum-norm solution is the subject of the following theorem, which both proves uniqueness and provides a recipe for the computation of the solution. The theorem is proven in an Appendix.

Theorem 2.1. *The minimum-norm least-squares solution to a linear system* $A\mathbf{x} = \mathbf{b}$*, that is, the shortest vector* $\mathbf{x}$ *that achieves the*

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\| \ ,$$

is unique, and is given by

$$\hat{\mathbf{x}} = V\Sigma^{\dagger}U^{T}\mathbf{b} \tag{2}$$

where

$$\Sigma^{\dagger} = \begin{bmatrix} 1/\sigma_1 & & & & & 0 & \cdots & 0 \\ & \ddots & & & & & & \\ & & 1/\sigma_r & & & \vdots & & \vdots \\ & & & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & 0 & 0 & \cdots & 0 \end{bmatrix}$$

The matrix

$$A^{\dagger} = V\Sigma^{\dagger}U^{T}$$

is called the *pseudoinverse* of A.

3 Least-Squares Solution of a Homogeneous Linear Systems

Theorem 2.1 works regardless of the value of the right-hand side vector $\mathbf{b}$. When $\mathbf{b} = \mathbf{0}$, that is, when the system is *homogeneous*, the solution is trivial: the minimum-norm solution to

$$A\mathbf{x} = \mathbf{0} \tag{3}$$

is

$$\mathbf{x} = \mathbf{0} \ ,$$

which happens to be an exact solution. Of course it is not necessarily the only one (any vector in the null space of A is also a solution, by definition), but it is obviously the one with the smallest norm.

Thus, $\mathbf{x} = \mathbf{0}$ is the minimum-norm solution to any homogeneous linear system. Although correct, this solution is not too interesting. In many applications, what is desired is a *nonzero* vector $\mathbf{x}$ that satisfies the system (3) as well as possible. Without any constraints on $\mathbf{x}$, we would fall back to $\mathbf{x} = \mathbf{0}$ again. For homogeneous linear systems, the meaning of a least-squares solution is therefore usually modified, once more, by imposing the constraint

$$\|\mathbf{x}\| = 1$$

on the solution. Unfortunately, the resulting constrained minimization problem does not necessarily admit a *unique* solution. The following theorem provides a recipe for finding this solution, and shows that there is in general a whole hypersphere of solutions.

Theorem 3.1. *Let*

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

be the singular value decomposition of A. Furthermore, let $\mathbf{v}_{n-k+1}, \ldots, \mathbf{v}_n$ be the k columns of V whose corresponding singular values are equal to the last singular value σ_n, that is, let k be the largest integer such that

$$\sigma_{n-k+1} = \ldots = \sigma_n \ .$$

Then, all vectors of the form

$$\mathbf{x} = \alpha_1 \mathbf{v}_{n-k+1} + \ldots + \alpha_k \mathbf{v}_n$$

with

$$\alpha_1^2 + \ldots + \alpha_k^2 = 1$$

are unit-norm least-squares solutions to the homogeneous linear system

$$A\mathbf{x} = \mathbf{0},$$

that is, they achieve the

$$\min_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \ .$$

Note: when σ_n is greater than zero the most common case is $k = 1$, since it is very unlikely that different singular values have *exactly* the same numerical value. When A is rank deficient, on the other case, it may often have more than one singular value equal to zero. In any event, if $k = 1$, then the minimum-norm solution is unique, $\mathbf{x} = \mathbf{v}_n$. If $k > 1$, then $\mathbf{x} = \mathbf{v}_n$ is still *a* unit-norm least-squares solution. In addition, in this case the theorem above shows how to express *all* solutions as a linear combination of the last k columns of V.

A Proofs

Theorem 2.1

The minimum-norm least-squares solution to a linear system $A\mathbf{x} = \mathbf{b}$, that is, the shortest vector $\mathbf{x}$ that achieves the

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\| ,$$

is unique, and is given by

$$\hat{\mathbf{x}} = V\Sigma^{\dagger}U^T\mathbf{b} \tag{4}$$

where

$$\Sigma^{\dagger} = \begin{bmatrix} 1/\sigma_1 & & & & & 0 & \cdots & 0 \\ & \ddots & & & & & & \\ & & 1/\sigma_r & & & \vdots & & \vdots \\ & & & 0 & & & & \\ & & & & \ddots & & & \\ & & & & 0 & 0 & \cdots & 0 \end{bmatrix}$$

Proof. The minimum-norm least-squares solution to

$$A\mathbf{x} = \mathbf{b}$$

is the shortest vector $\mathbf{x}$ that minimizes

$$\|A\mathbf{x} - \mathbf{b}\|$$

that is,

$$\|U\Sigma V^T\mathbf{x} - \mathbf{b}\| .$$

This can be written as

$$\|U(\Sigma V^T\mathbf{x} - U^T\mathbf{b})\| \tag{5}$$

because U is an orthogonal matrix, $UU^T = I$. But orthogonal matrices do not change the norm of vectors they are applied to, so that the last expression above equals

$$\|\Sigma V^T\mathbf{x} - U^T\mathbf{b}\|$$

or, with $\mathbf{y} = V^T\mathbf{x}$ and $\mathbf{c} = U^T\mathbf{b}$,

$$\|\Sigma\mathbf{y} - \mathbf{c}\| .$$

In order to find the solution to this minimization problem, let us spell out the last expression. We want to minimize the norm of the following vector:

$$\begin{bmatrix} \sigma_1 & 0 & & \cdots & & 0 \\ 0 & \ddots & & \cdots & & 0 \\ & & \sigma_r & & & \\ \vdots & & & 0 & & \vdots \\ & & & & \ddots & \\ 0 & & & & & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_r \\ y_{r+1} \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} c_1 \\ \vdots \\ c_r \\ c_{r+1} \\ \vdots \\ c_m \end{bmatrix} .$$

The last $m-r$ differences are of the form

$$\mathbf{0} - \begin{bmatrix} c_{r+1} \\ \vdots \\ c_m \end{bmatrix}$$

and do not depend on the unknown $\mathbf{y}$. In other words, there is nothing we can do about those differences: if some or all the c_i for $i = r+1, \ldots, m$ are nonzero, we will not be able to zero these differences, and each of them contributes a *residual* $|c_i|$ to the solution. Each of the first r differences, on the other hand, can be zeroed exactly by letting $y_i = c_i/\sigma_i$ for the i-th difference. In addition, in these first r differences, the last $n-r$ components of $\mathbf{y}$ are multiplied by zeros, so they have no effect on the solution. Thus, there is freedom in their choice. Since we look for the minimum-norm solution, that is, for the shortest vector $\mathbf{x}$, we also want the shortest $\mathbf{y}$, because $\mathbf{x}$ and $\mathbf{y}$ are related by an orthogonal transformation. We therefore set $y_{r+1} = \ldots = y_n = 0$. In summary, the desired $\mathbf{y}$ has the following components:

$$\begin{aligned} y_i &= \frac{c_i}{\sigma_i} \quad \text{for } i = 1, \ldots, r \\ y_i &= 0 \quad \text{for } i = r+1, \ldots, n \; . \end{aligned}$$

When written as a function of the vector $\mathbf{c}$, this is

$$\mathbf{y} = \Sigma^+ \mathbf{c} \; .$$

Notice that there is no other choice for $\mathbf{y}$, which is therefore unique: minimum residual forces the choice of $y_1, \ldots, y_r$, and minimum-norm solution forces the other entries of $\mathbf{y}$. Thus, the minimum-norm, least-squares solution to the original system is the unique vector

$$\mathbf{x} = V\mathbf{y} = V\Sigma^+\mathbf{c} = V\Sigma^+U^T\mathbf{b}$$

as promised. The residual, that is, the norm of $\|A\mathbf{x} - \mathbf{b}\|$ when $\mathbf{x}$ is the solution vector, is the norm of $\Sigma\mathbf{y} - \mathbf{c}$, since this vector is related to $A\mathbf{x} - \mathbf{b}$ by an orthogonal transformation (see equation (5)). In conclusion, the square of the residual is

$$\|A\mathbf{x} - \mathbf{b}\|^2 = \|\Sigma\mathbf{y} - \mathbf{c}\|^2 = \sum_{i=r+1}^{m} c_i^2 = \sum_{i=r+1}^{m} (\mathbf{u}_i^T \mathbf{b})^2$$

which is the projection of the right-hand side vector $\mathbf{b}$ onto the complement of the range of A.

Theorem 3.1

Let

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

be the singular value decomposition of A. Furthermore, let $\mathbf{v}_{n-k+1}, \ldots, \mathbf{v}_n$ be the k columns of V whose corresponding singular values are equal to the last singular value σ_n, that is, let k be the largest integer such that

$$\sigma_{n-k+1} = \ldots = \sigma_n \; .$$

Then, all vectors of the form

$$\mathbf{x} = \alpha_1 \mathbf{v}_{n-k+1} + \ldots + \alpha_k \mathbf{v}_n$$

with

$$\alpha_1^2 + \ldots + \alpha_k^2 = 1$$

are unit-norm least-squares solutions to the homogeneous linear system

$$A\mathbf{x} = \mathbf{0},$$

that is, they achieve the

$$\min_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \ .$$

Proof. The reasoning is similar to that for the previous theorem. The unit-norm least-squares solution to

$$A\mathbf{x} = \mathbf{0}$$

is the vector $\mathbf{x}$ with $\|\mathbf{x}\| = 1$ that minimizes

$$\|A\mathbf{x}\|$$

that is,

$$\|U\Sigma V^T\mathbf{x}\| \ .$$

Since orthogonal matrices do not change the norm of vectors they are applied to, this norm is the same as

$$\|\Sigma V^T\mathbf{x}\|$$

or, with $\mathbf{y} = V^T\mathbf{x}$,

$$\|\Sigma\mathbf{y}\| \ .$$

Since V is orthogonal, $\|\mathbf{x}\| = 1$ translates to $\|\mathbf{y}\| = 1$. We thus look for the unit-norm vector $\mathbf{y}$ that minimizes the norm (squared) of $\Sigma\mathbf{y}$, that is,

$$\sigma_1^2 y_1^2 + \ldots + \sigma_n^2 y_n^2 \ .$$

This is obviously achieved by concentrating all the (unit) mass of $\mathbf{y}$ where the σs are smallest, that is by letting

$$y_1 = \ldots = y_{n-k} = 0. \tag{6}$$

From $\mathbf{y} = V^T\mathbf{x}$ we obtain $\mathbf{x} = V\mathbf{y} = y_1\mathbf{v}_1 + \ldots + y_n\mathbf{v}_n$, so that equation (6) is equivalent to

$$\mathbf{x} = \alpha_1\mathbf{v}_{n-k+1} + \ldots + \alpha_k\mathbf{v}_n$$

with $\alpha_1 = y_{n-k+1}, \ldots, \alpha_k = y_n$, and the unit-norm constraint on $\mathbf{y}$ yields

$$\alpha_1^2 + \ldots + \alpha_k^2 = 1 \ .$$