Bounded Approximate Solutions of Linear Systems using SVD

Stephen Brooks

April 8, 2015

1 Definitions

The Singular Value Decomposition (SVD) of a complex matrix is conventionally $A = UV^*$, where $M^*$ denotes $M^T$. Here, $U$ and $V$ are unitary matrices with $U^{-1} = U^*$ and $\Sigma$ is diagonal with $\Sigma = \text{diag}[\sigma_n]$. For real matrices this is just $A = U\Sigma V^T$ and unitarity is equivalent to $U^{-1} = U^T$, i.e. orthogonality. In fact, $V^T$ is also orthogonal since $(V^T)^{-1} = (V^{-1})^{-1} = V = (V^T)^T$, which means the simpler definition $A = U\Sigma V$ can be used for the rest of this note.

2 Fundamental Problem

In control systems, one often uses a linear or locally-linear model to determine the required inputs. Suppose an input vector change $x \in X$ produces an output response $Ax \in Y$ that is meant to achieve some desired change $b \in Y$. The input and output spaces $X$ and $Y$ may have different dimensionalities and therefore $A$ can be a rectangular matrix. This means that an exact solution may not be possible, particularly if $\text{dim} \ Y > \text{dim} \ X$. Thus the ‘best’ solution can be formulated as the minimisation problem of finding $\arg \min |Ax - b|_Y$.

However, particularly in the case of ill-conditioned matrices, the exact solution may require unacceptably large control inputs. What is required practically is the best approximation that can be achieved while $x$ is not too large. This suggests casting the fundamental problem as

$$\arg \min_{|x|_X \leq r} |Ax - b|_Y$$

with $r > 0$ being chosen depending on how large a solution is acceptable. As $r \to \infty$, the value will eventually settle at the exact or optimum solution if one exists.

3 Solution using SVD

The SVD decomposition of $A$ gives

$$\arg \min_{|x|_X \leq r} |Ax - b|_Y = \arg \min_{|x|_X \leq r} |U\Sigma Vx - b|_Y.$$

Here, $A$ and $\Sigma$ are possibly-rectangular matrices mapping from $X$ to $Y$, $V$ is a square orthogonal matrix mapping $X$ to itself and $U$ is another mapping $Y$ to itself. Note that any orthogonal
matrix $U$ preserves the norm as $|Ux|^2 = x^TU^TUx = x^TU^{-1}Ux = x^T x = |x|^2$ so $|Ux| = |x|$ as norms are non-negative. In particular,

$$|x|_X = |Vx|_X \quad \text{and} \quad |USVx - b|_Y = |SVx - U^{-1}b|_Y,$$

where the second equality has multiplied by the unitary matrix $U^{-1}$. This means that

$$\text{arg} \min_{|x|_X \leq r} |Ax - b|_Y = \text{arg} \min_{|x|_X \leq r} |SVx - U^{-1}b|_Y.$$ 

Defining vectors $v = Vx$ and $u = U^{-1}b$ this becomes

$$\text{arg} \min_{|x|_X \leq r} |Ax - b|_Y = V^{-1} \text{arg} \min_{|v|_X \leq r} |Sv - u|_Y,$$

where the right-hand arg min is now understood to find the value of $v$, so the premultiplication for $x = V^{-1}v$ is required. The problem has now been simplified into one with a diagonal matrix instead of $A$.

### 3.1 Exact Minimum Solution

If the unrestricted arg min also satisfies $|x|_X \leq r$ then it is the solution. The unrestricted minimum is a fixed point of the norm expression squared:

$$0 = \frac{\partial}{\partial v_n} |\Sigma v - u|^2_Y = \frac{\partial}{\partial v_n} \sum_{i=1}^{\dim Y} (\Sigma v - u)_i^2 = \frac{\partial}{\partial v_n} \sum_{i=1}^{\dim Y} (\Sigma_i v_i - u_i)^2$$

$$= \frac{\partial}{\partial v_n} (\sigma_n v_n - u_n)^2 = \frac{\partial}{\partial v_n} (\sigma_n^2 v_n^2 - 2\sigma_n v_n u_n + u_n^2) = 2\sigma_n^2 v_n - 2\sigma_n u_n$$

$$\Leftrightarrow \sigma_n (\sigma_n v_n - u_n) = 0.$$ 

For each $n$, this is true if either $v_n = u_n/\sigma_n$ or $\sigma_n = 0$. In the latter case, the $\Sigma$ matrix does not range over the full dimensionality of $Y$ and any value of $v_n$ may be chosen because the minimum is non-unique. It is usually best to choose $v_n = 0$ in all such ambiguous cases, since this corresponds to the minimum with smallest $|v|_X = |x|_X$. There is also the case when $\dim Y < \dim X$, where the above equation reduces to $0 = 0$ for $n > \dim Y$, giving no constraint on $v_n$, which should be set to zero by the same argument. The exact minimum can be written explicitly as

$$x = V^{-1}[(U^{-1}b)_n/\sigma_n], \quad \text{where} \quad x/y = \begin{cases} x/y & \text{if } y \neq 0 \\ 0 & \text{otherwise} \end{cases}.$$ 

### 3.2 Constrained Minimum

The function $|\Sigma v - u|_Y$ does not have multiple disconnected local minima, so if the exact minimum with smallest norm found in the previous section still has $|x|_X > r$, the constrained minimum must have $|x|_X = r$ rather than being an interior point. The local gradient found in the previous section

$$\nabla_v |\Sigma v - u|^2_Y = 2[\sigma_n^2 v_n - \sigma_n u_n]$$

must be a scalar multiple of the position $v$ because otherwise it has some component parallel to the surface of the radius $r$ hypersphere and the value of the function can be reduced. The
gradient is expected to be negative with increasing \( r \), anti-parallel to \( v \), so for some \( \lambda > 0 \),

\[
\nabla v |\Sigma v - u|^2_Y = -2\lambda^2 v
\]

\[
\Leftrightarrow 2(\sigma_n^2 v_n - \sigma_n u_n) = -2\lambda^2 v_n
\]

\[
\Leftrightarrow (\sigma_n^2 + \lambda^2)v_n - \sigma_n u_n = 0
\]

\[
\Leftrightarrow v_n = \frac{\sigma_n u_n}{\sigma_n^2 + \lambda^2}.
\]

For the case where \( n > \dim Y \), the gradient of that component is zero as before and \( 0 = -2\lambda^2 v_n \), so \( v_n = 0 \). The constrained minimum can be written explicitly as

\[
x = V^{-1} \left[ \frac{\sigma_n (U^{-1}b)_n}{\sigma_n^2 + \lambda^2} \right], \quad \text{where we set} \quad (U^{-1}b)_n = 0 \quad \text{if} \ n > \dim Y.
\]

The norm of \( x \) decreases monotonically with \( \lambda \) because \( |x|_X = |v|_X \) and every element of \( v \) decreases in magnitude with increasing \( \lambda \). As \( \lambda \to 0 \) the constrained minimum tends towards the exact minimum. As \( \lambda \to \infty \), the constrained minimum tends towards \( 0 \) but if renormalised, the limit has \( v_n = \sigma_n u_n \), which is \(-\frac{1}{2}\) times the gradient of \( |\Sigma v - u|^2_Y \) at \( v = 0 \). Thus the large \( \lambda \) limit corresponds to a infinitesimal ‘steepest descent’ step.

The continuity and monotonicity of \( |x|_X = r(\lambda) \) ensures a value of \( \lambda \) can always be found for any value of \( r \) between 0 and the norm of the exact solution point. For example, a bisection search or root-finding algorithm can determine \( \lambda \) for a given \( r \), after first checking the exact solution point does not have norm less than \( r \).

### 3.3 Implementation Note

Using the orthogonal property of \( U \) and \( V \), entries \((U^{-1}b)_n\) should be calculated as the much faster equivalent \((U^Tb)_n\) and the premultiplication by \( V^{-1} \) should be implemented as \( V^T \). Once the SVD is calculated, nothing slower than matrix-vector multiplication is required.

### 4 Units

Elements of the vector spaces \( X \) and \( Y \) can be physical quantities with units \([X]\) and \([Y]\) respectively. By definition, \( A \) has units \([Y]/[X]\). In the SVD, the entries of \( U \) and \( V \) have no units as they map within the same space, leaving \( \Sigma \) and its entries \( \sigma_n \) with units \([Y]/[X]\). The parameter \( \lambda \) in the previous section was defined to also have units \([Y]/[X]\) but \( r \) has units \([X]\).