BNL-223624-2022-TECH

C-A/AP/671

Bounded Approximate Solutions of Linear Systems using SVD

S. Brooks

April 2015

Collider Accelerator Department

Brookhaven National Laboratory

U.S. Department of Energy

USDOE Office of Science (SC), Nuclear Physics (NP) (SC-26)

DISCLAIMER

Bounded Approximate Solutions of Linear Systems using SVD

Stephen Brooks

March 6, 2023

1 Definitions

The Singular Value Decomposition (SVD) of a complex matrix is conventionally $A = U\Sigma V^*$, where M^* denotes $\bar{M}^T$. Here, U and V are unitary matrices with $U^{-1} = U^*$ and Σ is diagonal with $\Sigma = \mathrm{diag}[\sigma_n]$. 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 $\mathbf{x} \in X$ produces an output reponse $A\mathbf{x} \in Y$ that is meant to achieve some desired change $\mathbf{b} \in Y$. 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 $\dim Y > \dim X$. Thus the 'best' solution can be formulated as the minimisation problem of finding $\arg\min |A\mathbf{x} - \mathbf{b}|_Y$.

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 $\mathbf{x}$ is not too large. This suggests casting the fundamental problem as

$$\arg \min_{|\mathbf{x}|_X \le r} |A\mathbf{x} - \mathbf{b}|_Y$$

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

$$\arg \min_{|\mathbf{x}|_X \le r} |A\mathbf{x} - \mathbf{b}|_Y = \arg \min_{|\mathbf{x}|_X \le r} |U\Sigma V\mathbf{x} - \mathbf{b}|_Y.$$

Here, A and Σ 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 $|U\mathbf{x}|^2 = \mathbf{x}^T U^T U \mathbf{x} = \mathbf{x}^T U^{-1} U \mathbf{x} = \mathbf{x}^T \mathbf{x} = |\mathbf{x}|^2$ so $|U\mathbf{x}| = |\mathbf{x}|$ as norms are non-negative. In particular,

$$|\mathbf{x}|_X = |V\mathbf{x}|_X \qquad \text{and} \qquad |U\Sigma V\mathbf{x} - \mathbf{b}|_Y = |\Sigma V\mathbf{x} - U^{-1}\mathbf{b}|_Y,$$

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

$$\arg\min_{|\mathbf{x}|_X \le r} |A\mathbf{x} - \mathbf{b}|_Y = \arg\min_{|V\mathbf{x}|_X \le r} |\Sigma V\mathbf{x} - U^{-1}\mathbf{b}|_Y.$$

Defining vectors $\mathbf{v} = V\mathbf{x}$ and $\mathbf{u} = U^{-1}\mathbf{b}$ this becomes

$$\arg\min_{|\mathbf{x}|_X \le r} |A\mathbf{x} - \mathbf{b}|_Y = V^{-1} \arg\min_{|\mathbf{v}|_X \le r} |\Sigma\mathbf{v} - \mathbf{u}|_Y,$$

where the right-hand $\arg\min$ is now understood to find the value of $\mathbf{v}$, so the premultiplication for $\mathbf{x} = V^{-1}\mathbf{v}$ 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 $|\mathbf{x}|_X \le r$ then it is the solution. The unrestricted minimum is a fixed point of the norm expression squared:

$$\begin{aligned}
0 &= \frac{\partial}{\partial v_n} |\Sigma\mathbf{v} - \mathbf{u}|_Y^2 = \frac{\partial}{\partial v_n} \sum_{i=1}^{\dim Y} (\Sigma\mathbf{v} - \mathbf{u})_i^2 = \frac{\partial}{\partial v_n} \sum_{i=1}^{\dim Y} (1_{i \le \dim X} \sigma_i v_i - u_i)^2 \\
&= \frac{\partial}{\partial v_n} (\sigma_n v_n - u_n)^2 = \frac{\partial}{\partial v_n} (\sigma_n^2 v_n^2 - 2\sigma_n v_n u_n + u_n^2) = 2\sigma_n^2 v_n - 2\sigma_n u_n \\
&\Leftrightarrow \quad \sigma_n(\sigma_n v_n - u_n) = 0.
\end{aligned}$$

For each n, this is true if either $v_n = u_n/\sigma_n$ or $\sigma_n = 0$. In the latter case, the Σ 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 $|\mathbf{v}|_X = |\mathbf{x}|_X$. 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

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

3.2 Constrained Minimum

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

$$\nabla_{\mathbf{v}} |\Sigma\mathbf{v} - \mathbf{u}|_Y^2 = 2[\sigma_n^2 v_n - \sigma_n u_n]$$

must be a scalar multiple of the position $\mathbf{v}$ 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 $\mathbf{v}$, so for some $\lambda > 0$,

$$\begin{array}{rrcl} & \nabla_{\mathbf{v}}|\Sigma\mathbf{v}-\mathbf{u}|_Y^2 & = & -2\lambda^2\mathbf{v} \\ \Leftrightarrow & 2(\sigma_n^2 v_n - \sigma_n u_n) & = & -2\lambda^2 v_n \\ \Leftrightarrow & (\sigma_n^2+\lambda^2)v_n - \sigma_n u_n & = & 0 \\ \Leftrightarrow & v_n & = & \dfrac{\sigma_n u_n}{\sigma_n^2+\lambda^2}. \end{array}$$

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

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

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

The continuity and monotonicity of $|\mathbf{x}|_X = r(\lambda)$ ensures a value of λ 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 λ 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 $(U^{-1}\mathbf{b})_n$ should be calculated as the much faster equivalent $(U^T\mathbf{b})_n$ 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 Σ and its entries σ_n with units $[Y]/[X]$. The parameter λ in the previous section was defined to also have units $[Y]/[X]$ but r has units $[X]$.

5 Identity with the Levenberg–Marquardt Algorithm

The Levenberg–Marquardt algorithm involves a 'damped' least squares step, which for a Jacobian matrix J involves solving

$$(J^TJ + \lambda_{LM}I)\mathbf{x} = J^T\mathbf{b},$$

where $\lambda_{LM} \geq 0$ is called the damping factor. If the Jacobian is decomposed via SVD as $J = U\Sigma V$, this becomes

$$(V^T\Sigma U^TU\Sigma V + \lambda_{LM}I)\mathbf{x} = V^T\Sigma U^T\mathbf{b}$$

and noting that $U^T U = I$ by orthogonality of U,

$$(V^T \Sigma^2 V + \lambda_{LM} I)\mathbf{x} = V^T \Sigma U^T \mathbf{b}.$$

Pre-multipliying both sides by V and using its orthogonality $VV^T = I$ gives

$$\begin{aligned} (\Sigma^2 V + \lambda_{LM} V)\mathbf{x} &= \Sigma U^T \mathbf{b} \\ \Rightarrow \quad (\Sigma^2 + \lambda_{LM}) V\mathbf{x} &= \Sigma U^T \mathbf{b}. \end{aligned}$$

This is starting to look vaguely familiar. Inverting the left-hand side to give an expression for $\mathbf{x}$ yields

$$\begin{aligned} \mathbf{x} &= V^{-1}(\Sigma^2 + \lambda_{LM} I)^{-1} \Sigma U^T \mathbf{b} \\ &= V^{-1}(\Sigma^2 + \lambda_{LM} I)^{-1} \Sigma U^{-1} \mathbf{b}. \end{aligned}$$

Comparing this to the constrained minimum formula with parameter λ from a previous section:

$$\mathbf{x} = V^{-1} \left[\frac{\sigma_n (U^{-1}\mathbf{b})_n}{\sigma_n^2 + \lambda^2} \right]$$

and noting that $\Sigma = \text{diag}[\sigma_n]$ reveals that these are the same formulae if $\lambda_{LM} = \lambda^2$.