Total Least Squares

CEE 629. System Identification

Duke University, Fall 2017

1 Problem Statement

Given a set of m data coordinates, $\{(x_1, y_1), \cdots, (x_m, y_m)\}$, a model to the data, $\hat{\boldsymbol{y}}(x; \boldsymbol{a})$ that is linear in n parameters $(a_1, \cdots, a_n)$, $\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{a}$, $(m > n)$, find the parameters to the model that 'best' satisfies the approximation,

$$\boldsymbol{y} \approx \boldsymbol{X}\boldsymbol{a} \ . \tag{1}$$

Since there are more equations than unknowns $(m > n)$, this is an overdetermined set of equations. If the measurements of the independent variables x_i are known precisely, then the only difference between the model, $\hat{y}(x_i; \boldsymbol{a})$, and the measured data y_i, must be due to measurement error in y_i and natural variability in the data that the model is unable to reflect. In such cases, the 'best' approximation minimizes a scalar-valued norm of the difference between the data points y_i and the model $\hat{y}(x_i; \boldsymbol{a})$. In such a case the approximation of equation (1) can be made an equality by incorporating the errors in $\boldsymbol{y}$,

$$\boldsymbol{y} + \tilde{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{a} \ , \tag{2}$$

where the vector $\tilde{\boldsymbol{y}}$ accounts for the differences between the measurements $\boldsymbol{y}$ and the model $\hat{\boldsymbol{y}}$. The model parameters $\boldsymbol{a}$ that minimize $\frac{1}{2}\sum \tilde{y}_i^2$ is called the *ordinary least squares* solution.

If, on the other hand, both the dependent variable measurements, y_i , and independent variable measurements, x_i, are measured with error, then the 'best' model minimizes a scalar-valued norm of both the difference between the measured and modeled independent variables, x_i, and dependent variables, y_i. In such a case the approximation of equation (1) can be made an equality by incorporating the errors in both $\boldsymbol{y}$ and $\boldsymbol{X}$,

$$\boldsymbol{y} + \tilde{\boldsymbol{y}} = \left[\boldsymbol{X} + \tilde{\boldsymbol{X}}\right] \boldsymbol{a} \tag{3}$$

Procedures for fitting a model to data that minimizes errors in both the dependent and independent variables are called *total least squares* methods. The problem, then, is to find the smallest perturbations $[\tilde{\boldsymbol{X}} \ \tilde{\boldsymbol{y}}]$ to the measured independent and dependent variables that satisfy the perturbed equation (3). This may be formally stated as

$$\min \left|\left| \left[\tilde{\boldsymbol{X}} \ \tilde{\boldsymbol{y}}\right] \right|\right|_{\mathsf{F}}^2 \quad \text{such that} \quad \boldsymbol{y} + \tilde{\boldsymbol{y}} = \left[\boldsymbol{X} + \tilde{\boldsymbol{X}}\right] \boldsymbol{a} \ , \tag{4}$$

where $[\tilde{\boldsymbol{X}} \ \tilde{\boldsymbol{y}}]$ is a matrix made from the column-wise concatenation of the matrix $\tilde{\boldsymbol{X}}$ with the vector $\tilde{\boldsymbol{y}}$, and $||[\boldsymbol{Z}]||_{\mathsf{F}}$ is the *Frobenius* norm of the matrix $\boldsymbol{Z}$,

$$||\boldsymbol{Z}||_{\mathsf{F}}^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} Z_{ij}^2 = \text{trace}\left(\boldsymbol{Z}^{\mathsf{T}}\boldsymbol{Z}\right) = \sum_{i=1}^{n} \sigma_i^2 \ , \tag{5}$$

where σ_i is the i-th singular value of the matrix $\boldsymbol{Z}$. So,

$$\left|\left|\left[\tilde{\boldsymbol{X}}\ \tilde{\boldsymbol{y}}\right]\right|\right|_{\mathsf{F}}^2 = \left|\left|\left[\begin{array}{cccc} | & & | & | \\ \tilde{\boldsymbol{x}}_1 & \cdots & \tilde{\boldsymbol{x}}_n & \tilde{\boldsymbol{y}} \\ | & & | & | \end{array}\right]\right|\right|_{\mathsf{F}}^2 = ||\tilde{\boldsymbol{x}}_1||_2^2 + \cdots + ||\tilde{\boldsymbol{x}}_n||_2^2 + ||\tilde{\boldsymbol{y}}||_2^2\ . \tag{6}$$

This minimization implies that the uncertainties in $\boldsymbol{y}$ and in $\boldsymbol{X}$ have comparable numerical magnitudes, in the units in which $\boldsymbol{y}$ and $\boldsymbol{X}$ are measured. One way to account for non-comparable levels of uncertainty would be to minimize $||[\tilde{\boldsymbol{X}}\ \gamma\tilde{\boldsymbol{y}}]||_{\mathsf{F}}^2$, where the scalar constant γ accounts for the relative errors in the measurements of $\boldsymbol{X}$ and $\boldsymbol{y}$. Alternatively, each column of $\boldsymbol{X}$ may be scaled individually.

2 Singular Value Decomposition

The singular value decomposition of an $m\times n$ matrix $\boldsymbol{X}$ is the product of three matrices,

$$\boldsymbol{X} = \boldsymbol{U\Sigma V}^{\mathsf{T}} \tag{7}$$

$$\boldsymbol{X} = \left[\begin{array}{cccc} | & | & & | \\ \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_m \\ | & | & & | \end{array}\right]_{m\times m} \left[\begin{array}{cccc} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \\ & & & \\ & & & \end{array}\right]_{m\times n} \left[\begin{array}{ccc} - & \boldsymbol{v}_1^{\mathsf{T}} & - \\ - & \boldsymbol{v}_2^{\mathsf{T}} & - \\ & \vdots & \\ - & \boldsymbol{v}_n^{\mathsf{T}} & - \end{array}\right]_{n\times n} \tag{8}$$

The matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthonormal, $\boldsymbol{U}^{\mathsf{T}}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^{\mathsf{T}} = \boldsymbol{I}_m$ and $\boldsymbol{V}^{\mathsf{T}}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{V}^{\mathsf{T}} = \boldsymbol{I}_n$. The first n rows of $\boldsymbol{\Sigma}$ are diagonal, with *singular values* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.

The action of the matrix $\boldsymbol{X}$ on the vector $\boldsymbol{a}$ to produce a vector $\hat{y}$, can be interpreted as the product of the actions of $\boldsymbol{V}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{U}$. First, the matrix $\boldsymbol{V}$ rotates $\boldsymbol{a}$ in the n-dimensional space of model parameters. Second, the matrix $\boldsymbol{\Sigma}$ scales the n components of the rotated vector $\boldsymbol{V}^{\mathsf{T}}\boldsymbol{a}$ in the rotated coordinate system. Last, the matrix $\boldsymbol{U}$ rotates the scaled vector $\boldsymbol{\Sigma V}^{\mathsf{T}}\boldsymbol{a}$ into the m dimensional coordinate system of data points. See the animation in [11].

Since the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthonormal, each column vector of $\boldsymbol{U}$ and $\boldsymbol{V}$ has unit length. The numerical values of the outer product of any pair of columns of $\boldsymbol{U}$ and $\boldsymbol{V}$, $[\boldsymbol{u}_i\boldsymbol{v}_i^{\mathsf{T}}]$ will have values of order 1. Also, since every row of the $m\times n$ matrix $[\boldsymbol{u}_i\boldsymbol{v}_i^{\mathsf{T}}]$ will be proportional to $\boldsymbol{v}_i^{\mathsf{T}}$, each matrix $[\boldsymbol{u}_i\boldsymbol{v}_i^{\mathsf{T}}]$ has a rank of one. Since $\boldsymbol{\Sigma}$ is diagonal up to the first n rows, the product $\boldsymbol{U\Sigma V}^{\mathsf{T}}$ can be written as a sum of rank-one matrices,

$$\boldsymbol{X} = \sum_{i=1}^{n} \sigma_i \left[\boldsymbol{u}_i\boldsymbol{v}_i^{\mathsf{T}}\right]. \tag{9}$$

Each term in the sum increases the rank of $\boldsymbol{X}$ by one. Each matrix $[\boldsymbol{u}_i\boldsymbol{v}_i^{\mathsf{T}}]$ has numerically about the same size. Since the singular values σ_i are ordered in decreasing numerical order, the term $\sigma_1[\boldsymbol{u}_1\boldsymbol{v}_1^{\mathsf{T}}]$ contributes the most to the sum, the term $\sigma_2[\boldsymbol{u}_2\boldsymbol{v}_2^{\mathsf{T}}]$ contributes a little less, and so on. The term $\sigma_n[\boldsymbol{u}_n \boldsymbol{v}_n^\mathsf{T}]$ contributes only negligibly. In many problems associated with the fitting of models to data, the spectrum of singular values has a sharp precipice, such that,

$$\sigma_1 \geq \sigma_2 \geq \cdots \sigma_{\hat{n}} \gg \sigma_{\hat{n}+1} \geq \cdots \geq \sigma_n \geq 0 \ .$$

In such cases the first $\hat{n}$ terms in the sum are numerically meaningful, and the terms $\hat{n}+1$ to n represent aspects of the model that do not contribute significantly.

The Eckart-Young theorem [3, 6] states that *the rank $\hat{n}$ matrix $\hat{Z}$ constructed from the first $\hat{n}$ terms in the SVD expansion of the rank n matrix $\boldsymbol{Z}$ minimizes $||[\hat{\boldsymbol{Z}} - \boldsymbol{Z}]||_\mathsf{F}^2$.*

The SVD of a matrix can be used to solve an over-determined set of equations in an ordinary least-squares sense. The orthonormality of $\boldsymbol{U}$ and $\boldsymbol{V}$, and the diagonallity of $\boldsymbol{\Sigma}$ makes this easy.

$$\begin{aligned}
\boldsymbol{y} &= \boldsymbol{X}\ \boldsymbol{a} \\
\boldsymbol{y} &= \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\mathsf{T}\ \boldsymbol{a} \\
\boldsymbol{U}^\mathsf{T}\ \boldsymbol{y} &= \boldsymbol{U}^\mathsf{T}\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\mathsf{T}\ \boldsymbol{a} = \boldsymbol{\Sigma}\boldsymbol{V}^\mathsf{T}\ \boldsymbol{a} \\
\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^\mathsf{T}\ \boldsymbol{y} &= \boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}\boldsymbol{V}^\mathsf{T}\ \boldsymbol{a} = \boldsymbol{V}^\mathsf{T}\ \boldsymbol{a} \\
\boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^\mathsf{T}\ \boldsymbol{y} &= \boldsymbol{V}\boldsymbol{V}^\mathsf{T}\ \boldsymbol{a} = \boldsymbol{a} \\
\boldsymbol{X}^+\ \boldsymbol{y} &= \boldsymbol{a} \qquad (10)
\end{aligned}$$

where $\boldsymbol{\Sigma}^{-1}$ is an $n \times n$ matrix with diagonal elements $1/\sigma_i$ and $\boldsymbol{X}^+ = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^\mathsf{T}$ is called the right Moore-Penrose generalized inverse of $\boldsymbol{X}$.

The generalized inverse, $\boldsymbol{X}^+ = \boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^\mathsf{T}$, can also be expanded into a sum

$$\boldsymbol{X}^+ = \sum_{i=1}^{n} \frac{1}{\sigma_i}\left[\boldsymbol{v}_i \boldsymbol{u}_i^\mathsf{T}\right] . \qquad (11)$$

where now the '$i = 1$' term, $\sigma_1^{-1}[\boldsymbol{v}_1\boldsymbol{u}_1^\mathsf{T}]$, contributes the least to the sum and the '$i = n$' $\sigma_n^{-1}[\boldsymbol{v}_n\boldsymbol{u}_n^\mathsf{T}]$, contributes the most. For problems in which the spectrum of singular values has a sharp precipice, at $i = \hat{n}$, truncating the series expansion of $\boldsymbol{X}^+$ at $i = \hat{n}$ eliminates the propagation of measurement errors in $\boldsymbol{y}$ to un-important aspects of the model parameters.

3 Singular Value Decomposition and Total Least Squares

Singular value decomposition can be used to find a unique solution to total least squares problems. The constraint equation (3) to the minimization problem (4) can be written,

$$\left[\boldsymbol{X} + \tilde{\boldsymbol{X}}\ ,\ \boldsymbol{y} + \tilde{\boldsymbol{y}}\right] \left[\begin{array}{c} \boldsymbol{a} \\ -1 \end{array}\right] = \boldsymbol{0}_{m \times 1}. \qquad (12)$$

The vector $[\boldsymbol{a}^\mathsf{T}\ ,\ -1]^\mathsf{T}$ lies in the null space of of the matrix $[\boldsymbol{X} + \tilde{\boldsymbol{X}}\ ,\ \boldsymbol{y} + \tilde{\boldsymbol{y}}]$.

Note that as long as the data vector $\boldsymbol{y}$ does not lie entirely in the sub-space spanned by the matrix $\boldsymbol{X}$, the concatenated matrix $[\boldsymbol{X}\ \boldsymbol{y}]$ has rank $n+1$. That is, the $n+1$ columns of $[\boldsymbol{X}\ \boldsymbol{y}]$ are linearly independent. The $n+1$, m-dimensional columns of the matrix $[\boldsymbol{X}\ \boldsymbol{y}]$ span the n dimensional space spanned by $\boldsymbol{X}$ and a component that is normal to the subspace spanned by $\boldsymbol{X}$.

For the solution $\boldsymbol{a}$ to be unique, the matrix $[\boldsymbol{X}+\tilde{\boldsymbol{X}}\ ,\ \boldsymbol{y}+\tilde{\boldsymbol{y}}]$ must have exactly n linearly independent columns. Since this matrix has $n+1$ columns in all, it must be rank-deficient by 1. Therefore, the goal of solving the minimization problem (4) can be restated as the goal of finding the "smallest" matrix $[\tilde{\boldsymbol{X}}\ \tilde{\boldsymbol{y}}]$ that changes $[\boldsymbol{X}\ \boldsymbol{y}]$ with rank $n+1$ to $[\boldsymbol{X}\ \boldsymbol{y}]+[\tilde{\boldsymbol{X}}\ \tilde{\boldsymbol{y}}]$ with rank n. The Eckart-Young Theorem provides the means to do so, by defining $[[\boldsymbol{X}\ \boldsymbol{y}]+[\tilde{\boldsymbol{X}}\ \tilde{\boldsymbol{y}}]]$ as the "best" rank-n approximation to $[\boldsymbol{X}\ \boldsymbol{y}]$. Dropping the last (smallest) singular value of $[\boldsymbol{X}\ \boldsymbol{y}]$ eliminates the least amount of information from the data and ensures a unique solution (assuming σ_{n+1} is not very close to σ_n).

The SVD of $[\boldsymbol{X}\ \boldsymbol{y}]$ can be partitioned

$$[\boldsymbol{X}\ \boldsymbol{y}] = [\boldsymbol{U}_p\ \boldsymbol{u}_q] \begin{bmatrix} \boldsymbol{\Sigma}_p & \\ & \sigma_q \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_{pp} & \boldsymbol{v}_{pq} \\ \boldsymbol{v}_{qp} & v_{qq} \end{bmatrix}^{\mathsf{T}} \tag{13}$$

where the matrix $\boldsymbol{U}_p$ has n columns, $\boldsymbol{u}_q$ is a column vector, $\boldsymbol{\Sigma}_p$ is diagonal with the n largest singular values, σ_q is the smallest singular value, $\boldsymbol{V}_{pp}$ is $n \times n$, and v_{qq} is a scalar. Note that in this representation of the SVD, $[\boldsymbol{U}_p\ \boldsymbol{u}_q]$ is of dimension $m \times (n+1)$, and the matrix of singular values is square.

Post-multiplying both sides of the SVD of $[\boldsymbol{X}\ \boldsymbol{y}]$ by $\boldsymbol{V}$, and equating just the last columns of the products,

$$\begin{aligned}
[\boldsymbol{X}\ \boldsymbol{y}] \begin{bmatrix} \boldsymbol{V}_{pp} & \boldsymbol{v}_{pq} \\ \boldsymbol{v}_{qp} & v_{qq} \end{bmatrix} &= [\boldsymbol{U}_p\ \boldsymbol{u}_q] \begin{bmatrix} \boldsymbol{\Sigma}_p & \\ & \sigma_q \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_{pp} & \boldsymbol{v}_{pq} \\ \boldsymbol{v}_{qp} & v_{qq} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \boldsymbol{V}_{pp} & \boldsymbol{v}_{pq} \\ \boldsymbol{v}_{qp} & v_{qq} \end{bmatrix} \\
[\boldsymbol{X}\ \boldsymbol{y}] \begin{bmatrix} \boldsymbol{V}_{pp} & \boldsymbol{v}_{pq} \\ \boldsymbol{v}_{qp} & v_{qq} \end{bmatrix} &= [\boldsymbol{U}_p\ \boldsymbol{u}_q] \begin{bmatrix} \boldsymbol{\Sigma}_p & \\ & \sigma_q \end{bmatrix} \\
[\boldsymbol{X}\ \boldsymbol{y}] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} &= \boldsymbol{u}_q \sigma_q
\end{aligned} \tag{14}$$

Now, *defining* $[\boldsymbol{X}\ \boldsymbol{y}]+[\tilde{\boldsymbol{X}}\ \tilde{\boldsymbol{y}}]$ to be the closest rank-n approximation to $[\boldsymbol{X}\ \boldsymbol{y}]$ in the sense of the Frobenius norm, $[\boldsymbol{X}\ \boldsymbol{y}]+[\tilde{\boldsymbol{X}}\ \tilde{y}]$ has the same singular vectors as $[\boldsymbol{X}\ \boldsymbol{y}]$ but only the n positive singular values contained in $\boldsymbol{\Sigma}_p$, [3, 6],

$$\left[\boldsymbol{X}+\tilde{\boldsymbol{X}}\ ,\ \boldsymbol{y}+\tilde{\boldsymbol{y}}\right] = [\boldsymbol{U}_p\ \boldsymbol{u}_q] \begin{bmatrix} \boldsymbol{\Sigma}_p & \\ & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_{pp} & \boldsymbol{v}_{pq} \\ \boldsymbol{v}_{qp} & v_{qq} \end{bmatrix}^{\mathsf{T}} . \tag{15}$$

Substituting equations (15) and (13) into the equality $[\tilde{\boldsymbol{X}}\ \tilde{\boldsymbol{y}}] = [\boldsymbol{X}\ \boldsymbol{y}]+[\tilde{\boldsymbol{X}}\ \tilde{\boldsymbol{y}}]-[\boldsymbol{X}\ \boldsymbol{y}]$, and then making use of equation (14),

$$
\begin{aligned}
\left[\tilde{\boldsymbol{X}} \ \tilde{\boldsymbol{y}}\right] &= -\left[\boldsymbol{U}_p \ \boldsymbol{u}_q\right] \begin{bmatrix} \boldsymbol{0}_{n\times n} & \\ & \sigma_q \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_{pp} & \boldsymbol{v}_{pq} \\ \boldsymbol{v}_{qp} & v_{qq} \end{bmatrix}^{\mathsf{T}} \\
&= -\left[\boldsymbol{0}_{m\times n} \ \boldsymbol{u}_q \sigma_q\right] \begin{bmatrix} \boldsymbol{V}_{pp} & \boldsymbol{v}_{pq} \\ \boldsymbol{v}_{qp} & v_{qq} \end{bmatrix}^{\mathsf{T}} \\
&= -\boldsymbol{u}_q \sigma_q \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix}^{\mathsf{T}} \\
&= -\left[\boldsymbol{X} \ y\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix}^{\mathsf{T}}
\end{aligned}
\tag{16}
$$

So,

$$
\begin{aligned}
\left[\boldsymbol{X}+\tilde{\boldsymbol{X}} \ , \ \boldsymbol{y}+\tilde{\boldsymbol{y}}\right] &= \left[\boldsymbol{X} \ \boldsymbol{y}\right] - \left[\boldsymbol{X} \ \boldsymbol{y}\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix}^{\mathsf{T}} \\
\left[\boldsymbol{X}+\tilde{\boldsymbol{X}} \ , \ \boldsymbol{y}+\tilde{\boldsymbol{y}}\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} &= \left[\boldsymbol{X} \ \boldsymbol{y}\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} - \left[\boldsymbol{X} \ \boldsymbol{y}\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix}^{\mathsf{T}} \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} \\
\left[\boldsymbol{X}+\tilde{\boldsymbol{X}} \ , \ \boldsymbol{y}+\tilde{\boldsymbol{y}}\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} &= \left[\boldsymbol{X} \ \boldsymbol{y}\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} - \left[\boldsymbol{X} \ \boldsymbol{y}\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} \\
\left[\boldsymbol{X}+\tilde{\boldsymbol{X}} \ , \ \boldsymbol{y}+\tilde{\boldsymbol{y}}\right] \begin{bmatrix} \boldsymbol{v}_{pq} \\ v_{qq} \end{bmatrix} &= \boldsymbol{0}_{m\times 1}
\end{aligned}
\tag{17}
$$

Comparing equation (17) with equation (12) we arrive at the unique total least squares solution for the vector of model parameters, $\boldsymbol{a}$,

$$
\boldsymbol{a}^*_{\mathsf{tls}} = -\boldsymbol{v}_{pq} \, v_{qq}^{-1} \ , \tag{18}
$$

where the vector $\boldsymbol{v}_{pq}$ is first n elements of the $(n+1)$-th column of the matrix of right singular vectors, $\boldsymbol{V}$, of $[\boldsymbol{X} \ \boldsymbol{y}]$ and v_{qq} is the $n+1$ element of the $n+1$ column of $\boldsymbol{V}$. Finally, the "curve-fit" is provided by $\hat{y}_{\mathsf{tls}} = [\boldsymbol{X}+\tilde{\boldsymbol{X}}]\boldsymbol{a}^*_{\mathsf{tls}}$ and requires the parameters, $\boldsymbol{a}^*_{\mathsf{tls}}$, *as well as* the perturbation in the basis, $\tilde{\boldsymbol{X}}$. Parameter values obtained from TLS can not be compared directly to those from OLS because the TLS solution is in terms of a different basis. This last point complicates the application of total least squares to curve-fitting problems in which a parameterized functional form $\hat{y}(x;\boldsymbol{a})$ is ultimately desired. As will be seen in the following example, computing $\hat{\boldsymbol{y}}$ using $\boldsymbol{X}\boldsymbol{a}^*_{\mathsf{tls}}$ can give bizarre results.

4 Numerical Examples

4.1 Example 1

Consider fitting a function $\hat{y}(x; \boldsymbol{a}) = a_1 x + a_2 x^2$ through three data points. The experiment is carried out twice, producing data $\boldsymbol{y}_{\mathsf{a}}$ and $\boldsymbol{y}_{\mathsf{b}}$.

i	x_i	$y_{\mathsf{a}i}$	$y_{\mathsf{b}i}$
1	1	1	2
2	5	3	3
3	7	8	8

For this problem, the *design* matrix $\boldsymbol{X}$ is

$$\boldsymbol{X} = \left[\begin{array}{cc} 1 & 1 \\ 5 & 25 \\ 7 & 49 \end{array} \right] . \tag{19}$$

The two columns of $\boldsymbol{X}$ are linearly independent (but not mutually orthogonal), and each column of $\boldsymbol{X}$ lies in a three-dimensional space: the space of the m data values. The columns of $\boldsymbol{X}$ span a two-dimensional *subspace* of the three-dimensional space of data values. The matrix equation $\hat{\boldsymbol{y}} = \boldsymbol{X}\boldsymbol{a}$ is simply a linear combination of the two columns of $\boldsymbol{X}$ and the three-dimensional vector $\hat{\boldsymbol{y}}$ must lie in the two-dimensional subspace spanned by $\boldsymbol{X}$. In other words, the vector $\hat{\boldsymbol{y}}$ must conform to the model $a_1 x + a_2 x^2$.

The SVD of $\boldsymbol{X}$ is computed in Matlab using `[ U, S, V ] = svd(X);`

$$\boldsymbol{U} = \left[\begin{array}{rrr} 0.02 & 0.55 & -0.83 \\ 0.46 & 0.74 & 0.50 \\ 0.89 & -0.39 & -0.24 \end{array} \right] \quad \boldsymbol{\Sigma} = \left[\begin{array}{rr} 55.68 & 0 \\ 0 & 1.51 \\ 0 & 0 \end{array} \right] \quad \boldsymbol{V} = \left[\begin{array}{rr} 0.15 & 0.99 \\ 0.99 & -0.15 \end{array} \right]$$

The dyadic contributions are

$$\hat{\boldsymbol{X}}_1 = \sigma_1[\boldsymbol{u}_1 \boldsymbol{v}_1^{\mathsf{T}}] = \left[\begin{array}{rr} 0.17 & 1.13 \\ 3.90 & 25.17 \\ 7.58 & 48.91 \end{array} \right] \quad \hat{\boldsymbol{X}}_2 = \sigma_2[\boldsymbol{u}_2 \boldsymbol{v}_2^{\mathsf{T}}] = \left[\begin{array}{rr} 0.83 & -0.13 \\ 1.10 & -0.17 \\ -0.58 & 0.09 \end{array} \right]$$

The first dyad is close to $\boldsymbol{X}$, and the second dyad contributes about two percent (σ_2/σ_1).

In Matlab there are at least three ways to compute the pseudo-inverse:

```
X_pi_ls  = inv(X'*X)*X'                     % least squares pseudo-inverse
X_pi_svd =  V*inv(S(1:n,1:n))*U(:,1:n)'     % SVD pseudo-inverse   eq'n (9)
X_pi     = X \ eye(m)                       % matlab "\" . . . like svd
```

and for most problems they provide nearly identical results. The orthogonal projection matrix from the data to the model can be computed from any of these, for example,

```
  PI_proj = X * ( X \ eye(3) );                % projection matrix using "\"
```

The ordinary least squares fit to each data set results in parameters

```
 a_ols = [ -0.21914 ;  0.18856 ]     % DATA SET A
 a_ols = [  0.14298 ;  0.13279 ]     % DATA SET B
```

To compute the total least squares solution in MATLAB,

```
  [ U, S, V ] = svd([X y]);                    % compute the SVD
  a_tls = - V(1:n,1+n) / V(n+1,n+1)            % total least squares soln  (17)
```

Other properties of the total least squares solution can be computed as ...

```
  Xtyt = - [X y] * V(:,n+1) * V(:,n+1)';       % [ X-tilde  y-tilde ] eq'n (15)
  Xtyt_norm_{\sf F} = norm(Xtyt,'fro')               % Frobeneus norm of [Xt yt] (5)
  Xt = Xtyt(:,1:n);                            % just X-tilde
```

The complete total least squares solutions, $\hat{\boldsymbol{y}}_{\sf tls} = \left[\boldsymbol{X} + \tilde{\boldsymbol{X}}\right] \boldsymbol{a}^*_{\sf tls}$, are listed below and plots are shown in Figures 1 and 2

Data Set A:

$$\begin{bmatrix} 0.036593 \\ 3.183288 \\ 7.930877 \end{bmatrix} = \begin{bmatrix} 0.47089 & 1.23103 \\ 5.10066 & 24.95605 \\ 6.96204 & 49.01658 \end{bmatrix} \begin{bmatrix} -0.54921 \\ 0.23981 \end{bmatrix} \tag{20}$$

Data Set B:

$$\begin{bmatrix} 2.0783 \\ 3.1647 \\ 7.9136 \end{bmatrix} = \begin{bmatrix} 0.42928 & 1.07634 \\ 3.80013 & 25.16051 \\ 7.62973 & 48.91576 \end{bmatrix} \begin{bmatrix} 7.28481 \\ -0.97448 \end{bmatrix} \tag{21}$$

- First, note that $\left[\boldsymbol{X} + \tilde{\boldsymbol{X}}\right] \neq \boldsymbol{X}$. The TLS basis has changed, as would be expected.
- Second, note that while the OLS fit is in a basis $[x,\ x^2]$, in the TLS solution the second column of $\left[\boldsymbol{X} + \tilde{\boldsymbol{X}}\right]$ is *not* the square of the first column. This begs the question, "How can the TLS solution be used to obtain values of y at other values of x?" (??)
- Third, even though the bases $[\boldsymbol{X}+\tilde{\boldsymbol{X}}]$ are close to $\boldsymbol{X}$, the difference between $[\boldsymbol{X}+\tilde{\boldsymbol{X}}]\boldsymbol{a}^*_{\sf tls}$ and $\boldsymbol{X}\boldsymbol{a}^*_{\sf tls}$ can be huge, as shown in Figure 2.

Golub and Van Loan [5] have shown that the conditioning of total least squares problems is always worse than that of ordinary least squares. In this sense total least squares is a "*de*regularization" of the ordinary least squares problem, and that the difference between the ordinary least squares and total least squares solutions grows with $1/(\sigma_n - \sigma_{n+1})$.

4.2 Example 2

This example involves more data points and is designed to test the OLS and TLS solutions with a problem involving measurement errors in both x and y. The "truth" model here is

$$y = 0.1x + 0.1x^2$$

and the "exact" values of x are $[1, 2, \cdots, 10]^{\mathsf{T}}$. Measurement errors are simulated with normally distributed random variables having a mean of zero and a standard deviation of 0.5. Errors are added to samples of *both* x and y. An example of the OLS and TLS solutions is shown in Figure 3. For this problem $\tilde{\boldsymbol{X}}$ is less than one percent of $\boldsymbol{X}$, so the TLS solution does not perturb the basis significantly, and the solution $\boldsymbol{X}\boldsymbol{a}^*_{\mathsf{tls}}$ (red dots) is hard to distinguish from the OLS solution (blue line). In fact, the curve $\boldsymbol{X}\boldsymbol{a}^*_{\mathsf{tls}}$ seems closer to the OLS solution than $[\boldsymbol{X} + \tilde{\boldsymbol{X}}]\boldsymbol{a}^*_{\mathsf{tls}}$ (red +). This analysis is repeated 1000 times and the estimated OLS and TLS coefficients are plotted in Figure 4. The average and standard deviation of the OLS and TLS parameters is given in in Table 1.

Table 1. Parameter estimates from 50,000 OLS and TLS estimates.

	mean OLS	std.dev OLS	mean TLS	std.dev TLS
a^*_1	0.099688	0.102826	0.107509	0.112426
a^*_2	0.100048	0.012663	0.099123	0.013787

Table 2. Parameter covariance matrices from 50,000 OLS and TLS estimates.

V_{a^*}	COV of OLS estimates	COV of TLS estimates
$\begin{bmatrix} 0.00902 & -0.00107 \\ -0.00107 & 0.00013 \end{bmatrix}$	$\begin{bmatrix} 0.01057 & -0.00126 \\ -0.00126 & 0.00016 \end{bmatrix}$	$\begin{bmatrix} 0.01264 & -0.00151 \\ -0.00151 & 0.00019 \end{bmatrix}$

The parameter estimates from such sparse and noisy data have a coefficient of variation of about 100 percent for a_1 and 10 percent for a_2. Table 2 indicates that a positive change in a_1 is usually accompanied by a negative change in a_2. This makes sense from the specifics of the problem being solved, and is clearly illustrated in Figure 4. The eigenvalues of V_{a^*} are are on the order 10^{-5} and 10^{-2} and the associated (orthonormal) eigenvectors are $[\ -0.12 \quad -0.99\ ]^{\mathsf{T}}$ and $[\ -0.99 \quad 0.12\ ]^{\mathsf{T}}$, representing the lengths and directions of the principal axes of the ellipsoidal distribution of parameters, shown in Figure 4.

From these 1000 analyses it is hard to discern if the total least squares method is advantageous in estimating model parameters from data with errors in both the dependent and independent variables. But the model used ($\hat{y}(x; a) = a_1 x + a_2 x^2$) is not particularly challenging. The relative merits of these methods would be better explored on a less well-conditioned problem, but using methodology similar to that used in this example.

5 Summary

Total least squares methods involve minimum perturbations to the basis of the model to minimally perturb the data to lie within the perturbed basis. The solution minimizes the Frobenius norm of these perturbations and can be related to the "best" n-rank approximation to the $n + 1$ rank matrix of the bases and the data. The method finds application in the modeling of data in which errors appear in the independent and dependent variables. This document was written in an attempt to understand and clarify the method by presenting a derivation and examples. Additional examples could address the effects of relative weightings or scaling between the basis and the data, and application to high-dimensional or ill-conditioned systems.

References

[1] Austin, David, "We Recommend a Singular Value Decomposition," American Mathematical Society, Feature Column,

[2] Cadzow, James A., "Total Least Squares, Matrix Enhancement, and Signal Processing," *Digital Signal Processiong,* 4: 23-39, (1994).

[3] Eckart, C. and Young, G., "The approximation of one matrix by another of lower rank". *Psychometrika,* 1936; 1(3): 211-8. doi:10.1007/BF02288367.

[4] Golub, G.H. and Reinsch, C. "Singular Value Decomposition and Least Squares Solutions," *Numer. Math.* 1970; 14:403-420.

[5] Golub, G.H. and Van Loan, C.F. "An Analysis of the Total Least Squares Problem," *SIAM J. Numer. Analysis.* 1970; 17(6):883-983.

Golub, G.H., Hansen, P.C., and O'Leary, D.P., "Tikhonov Regularization and Total Least Squares," *SIAM J. Matrix Analysis.* 1999; 21(1):185-194.

[6] Johnson, R.M., "On a Theorem Stated by Eckart and Young," *Psychometrika,* 1963; 28(3): 259-263.

[7] Press et.al., *Numerical Recipes in C*, second edition, 1992 (section 2.6)

[8] Stewart, G. W., "On the Early History of the Singular Value Decomposition". *SIAM Review,* 1993; 35(4): 551-566. doi:10.1137/1035134.

[9] Emanuele Trucco and Alessandro Verri, *Introductory Techniques for 3-D Computer Vision,* Prentice Hall, 1998. (Appendix A.6)

[10] Todd Will, Introduction to the Singular Value Decomposition. UW-La Crosse

[11] Wikipedia, Singular Value Decomposition, 2013.

[12] Wikipedia, Total Least Squares, 2013.

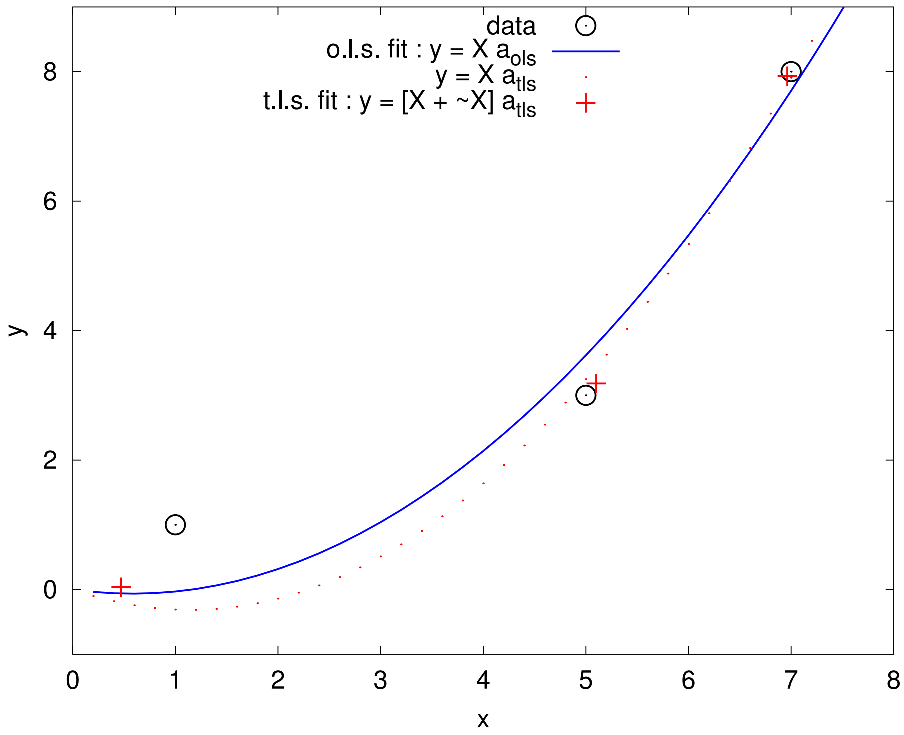

Figure 1. Data set "A": (o); OLS fit: (-); $\boldsymbol{X}\boldsymbol{a}^*_{\sf tls}$: (.); and $[\boldsymbol{X}+\tilde{\boldsymbol{X}}]\boldsymbol{a}^*_{\sf tls}$: (+)

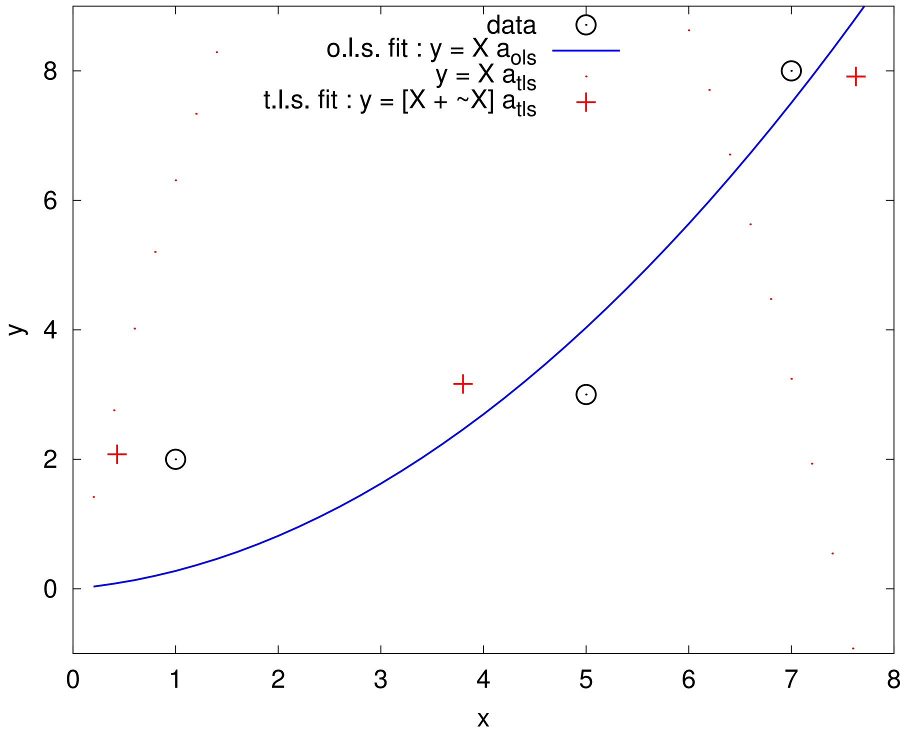

Figure 2. Data set "B": (o); OLS fit: (-); $\boldsymbol{X}\boldsymbol{a}^*_{\sf tls}$: (.); and $[\boldsymbol{X}+\tilde{\boldsymbol{X}}]\boldsymbol{a}^*_{\sf tls}$: (+)

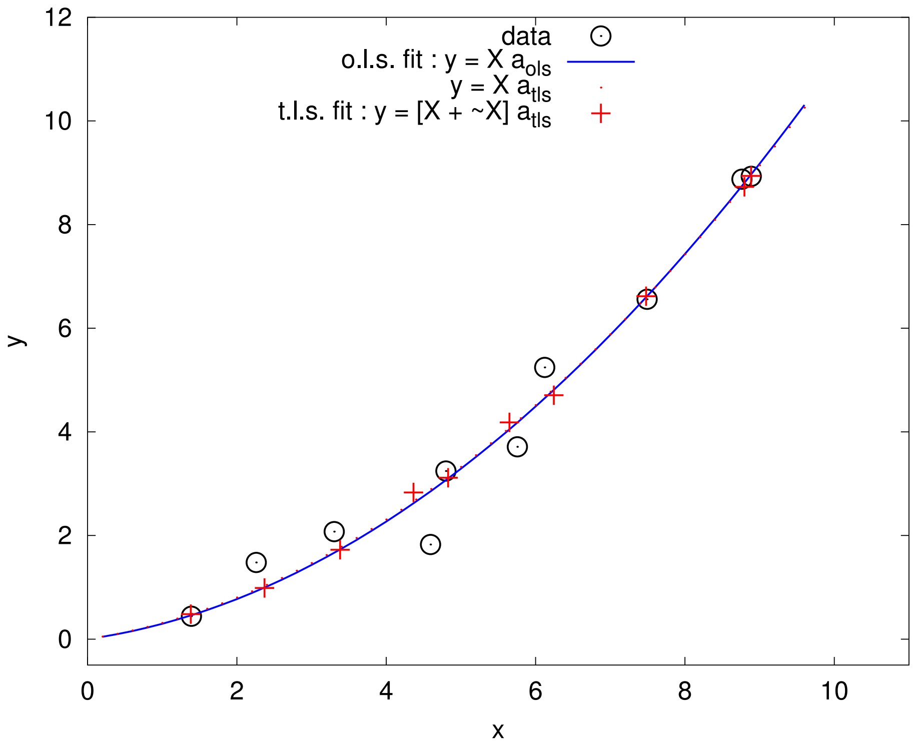

Figure 3. Data with errors in x and y: (o); OLS fit: (-); $\boldsymbol{X}\boldsymbol{a}^*_{\mathsf{tls}}$: (.); and $[\boldsymbol{X} + \tilde{\boldsymbol{X}}]\boldsymbol{a}^*_{\mathsf{tls}}$: (+)

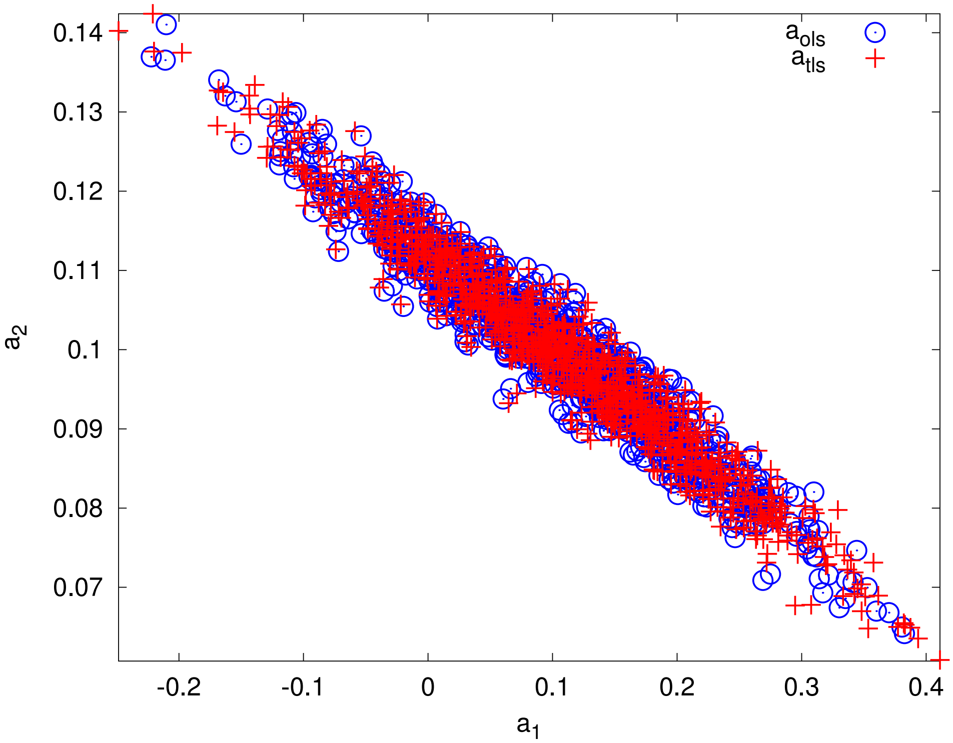

Figure 4. Propagation of measurement error to model parameters for OLS (o) and TLS (+)