Numerical Methods for Solving Least Squares Problems with Constraints

Gene H. Golub
Stanford University

The problems

Let A be a given $m \times n$ matrix of rank r and let $\mathbf{b}$ a given vector.

- **Linear least squares**:
  Find $\hat{\mathbf{x}}$ so that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\|_2 = \min\,.$$

- **Least squares with linear constraints**:
  Find $\hat{\mathbf{x}}$ so that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\|_2 = \min$$

  subject to

$$C^T\hat{\mathbf{x}} = \mathbf{0}.$$

- **Least squares with quadratic constraints**: Find $\hat{\mathbf{x}}$ so that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\|_2 = \min$$

  subject to

$$\|\hat{\mathbf{x}}\|_2^2 \leq \alpha^2.$$

- **Total least squares**:
  Find $\hat{\mathbf{x}}$, a matrix $\hat{E}$, and a residual $\hat{\mathbf{r}}$ so that

$$\left(\|\hat{E}\|_F^2 + \|\hat{\mathbf{r}}\|_2^2\right) = \min$$

  subject to

$$(A + \hat{E})\hat{\mathbf{x}} = \mathbf{b} + \hat{\mathbf{r}}.$$

- **Least squares with linear and quadratic constraints**: Find $\hat{\mathbf{x}}$ so that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\|_2 = \min$$

  subject to

$$C^T\hat{\mathbf{x}} = \mathbf{0} \text{ and } \|\hat{\mathbf{x}}\|_2^2 \leq \alpha^2.$$

Applications

- Statistical methods
- Image processing
- Data interpolation and surface fitting
- Geophysical problems
- ...

Method of solution depends upon

- Sparsity of matrix
- Size of the problem
- Accuracy required
- Application
- ...

Program

- Linear least squares
  - QR - decomposition
  - Singular systems
- Least squares with linear constraints
  - Lagrange multiplier
  - Augmented Lagrangian approach
    - GMRES applied to the KKT system
    - Uzawa algorithm
  - Weighting method
  - Direct method
- Least squares with quadratic constraints
  - Lagrange multipliers
    - The SVD/Newton approach
    - The quadratic eigenvalue approach
    - Approximating the secular equation

- Total least squares
  - Partial total least squares
  - Regularized total least squares
- Least squares with linear and quadratic constraints

1. Linear Least Squares

To solve the linear least squares problem accurately, we perform the following steps.

1. $$Q^T A = R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \begin{matrix} n \\ m-n \end{matrix}$$

   where R_1 is an upper triangular matrix.

2. $$Q^T \mathbf{b} = \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} \begin{matrix} n \\ m-n \end{matrix}$$

3. $$R_1 \hat{\mathbf{x}} = \mathbf{c}$$

This will yield a solution even when A is not of full rank.

The decomposition is performed via

1. Householder Transformations
2. Givens Rotations
3. Modified Gram-Schmidt Algorithm

We try to avoid using normal equations.

The QR factorization is useful for

- updating
- adding/deleting variables
- downdating.

Nevertheless the problem can be very ill - conditioned.

Theorem Suppose $\hat{\mathbf{x}}$, $\hat{\mathbf{r}}$, $\tilde{\mathbf{x}}$, and $\tilde{\mathbf{r}}$ satisfy

$$\begin{aligned} \|A\hat{\mathbf{x}} - \mathbf{b}\|_2 &= \min \\ \|(A + \delta A)\tilde{\mathbf{x}} - (\mathbf{b} + \delta\mathbf{b})\|_2 &= \min, \end{aligned}$$

$$\hat{\mathbf{r}} = \mathbf{b} - A\hat{\mathbf{x}}, \quad \tilde{\mathbf{r}} = (\mathbf{b} + \delta\mathbf{b}) - (A + \delta A)\tilde{\mathbf{x}},$$

where $m \geq n$ and $\mathbf{0} \neq \mathbf{b}$.

If

$$\epsilon = \max\left\{\frac{\|\delta A\|_2}{\|A\|_2}, \frac{\|\delta\mathbf{b}\|_2}{\|\mathbf{b}\|_2}\right\} < \frac{\sigma_n(A)}{\sigma_1(A)}$$

and

$$\sin(\theta) = \frac{\hat{\rho}}{\|\mathbf{b}\|_2} < 1$$

with $\hat{\rho} = \|A\hat{\mathbf{x}} - \mathbf{b}\|_2$, then

$$\frac{\|\tilde{\mathbf{x}} - \hat{\mathbf{x}}\|_2}{\|\hat{\mathbf{x}}\|_2} \leq \epsilon\left\{\frac{2\kappa_2(A)}{\cos(\theta)} + \tan(\theta)\kappa_2(A)^2\right\} + O(\epsilon^2)$$

$$\frac{\|\tilde{\mathbf{r}} - \hat{\mathbf{r}}\|_2}{\|\mathbf{b}\|_2} \leq \epsilon(1 + 2\kappa_2(A))\ \min(1, m-n) + O(\epsilon^2).$$

Singular Systems

To solve the linear least squares problem for matrices A which doesn't have full rank, perform the following two steps.

- Compute a complete orthogonal factorization

$$A = Q \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} Z^T$$

  where $Q^T Q = I_m$, $Z^T Z = I_n$, and R is an $r \times r$ upper triangular matrix.

- Compute the *pseudo-inverse* A^+ of A

$$A^+ = Z \begin{pmatrix} R^{-1} & 0 \\ 0 & 0 \end{pmatrix} Q^T$$

$$\begin{array}{llll} 1) & AA^+A = A & 3) & (AA^+)^T = (AA^+) \\ 2) & A^+AA^+ = A^+ & 4) & (A^+A)^T = (A^+A) \end{array}$$

Then $\hat{\mathbf{x}} = A^+\mathbf{b}$ is the solution of

$$\|\mathbf{b} - A\hat{\mathbf{x}}\|_2 = \min \quad \text{and} \quad \|\hat{\mathbf{x}}\|_2 = \min\,.$$

2. Least Squares with Linear Constraints

Consider

$$\|\mathbf{b} - A\mathbf{x}\|_2 = \min$$
$$\text{s.t. } C^T\mathbf{x} = \mathbf{0}$$

2.1 Lagrange multipliers

$$\psi(\mathbf{x};\boldsymbol{\lambda}) = \|\mathbf{b} - A\mathbf{x}\|_2^2 + 2\mathbf{x}^T C\boldsymbol{\lambda}$$

grad $\psi = \mathbf{0}$ when

$$\begin{aligned} A^T A\mathbf{x} + C\boldsymbol{\lambda} &= A^T\mathbf{b} \\ C^T\mathbf{x} &= \mathbf{0} \end{aligned}$$

or

$$\boxed{\begin{pmatrix} A^T A & C \\ C^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} A^T\mathbf{b} \\ \mathbf{0} \end{pmatrix}}$$

This system is known as the *KKT system.*

Direct method for the Lagrange multiplier approach

Let $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$ denote the solution of the unconstrained least squares problem. Then the first equation of the KKT system reads

$$\begin{aligned} \mathbf{x} &= (A^TA)^{-1}A^T\mathbf{b} - (A^TA)^{-1}C\boldsymbol{\lambda} \\ &= \hat{\mathbf{x}} - (A^TA)^{-1}C\boldsymbol{\lambda} \end{aligned}$$

which together with the second equation leads to

$$\begin{aligned} C^T(A^TA)^{-1}C\boldsymbol{\lambda} &= C^T\hat{\mathbf{x}} \\ C^TR_1^{-1}(R_1^T)^{-1}C\boldsymbol{\lambda} &= C^T\hat{\mathbf{x}}. \end{aligned}$$

The QR factorization of $(R_1^T)^{-1}C$ can be efficiently used for this solution.

2.2 Augmented Lagrangian approach

Original KKT system

$$\begin{aligned} A^T A\mathbf{x} + C\boldsymbol{\lambda} &= A^T\mathbf{b} \\ C^T\mathbf{x} &= \mathbf{0}. \end{aligned}$$

Since $CWC^T\mathbf{x} = \mathbf{0}$, can rewrite system as

$$\begin{aligned} (A^T A + CWC^T)\mathbf{x} + C\boldsymbol{\lambda} &= A^T\mathbf{b} \\ C^T\mathbf{x} &= \mathbf{0} \end{aligned}$$

Useful when $A^T A$ is (almost) singular.

Picking the right $\mathbf{W}$:

- Sparsity considerations (e.g. use only certain columns of C).
- Obtain positive definiteness of the (1,1) block.
- Scaling/Balancing. For example set $W = \gamma I$, $\gamma = \frac{\|A^T A\|}{\|C\|^2}$.

Estimating the condition number

Theorem. Suppose that A is a $m \times n$ matrix, C is a full column rank $n \times p$ matrix ($p \leq n$), and W is a $p \times p$ matrix. Define

$$\mathcal{A}(W) := \begin{pmatrix} A^T A + CWC^T & C \\ C^T & 0 \end{pmatrix}.$$

Then for any $W \neq 0$ such that $\mathcal{A}(W)$ is nonsingular, the following holds:

$$\mathcal{A}^{-1}(W) = \mathcal{A}^{-1} - \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}, \quad \text{where } \mathcal{A} \equiv \mathcal{A}(0).$$

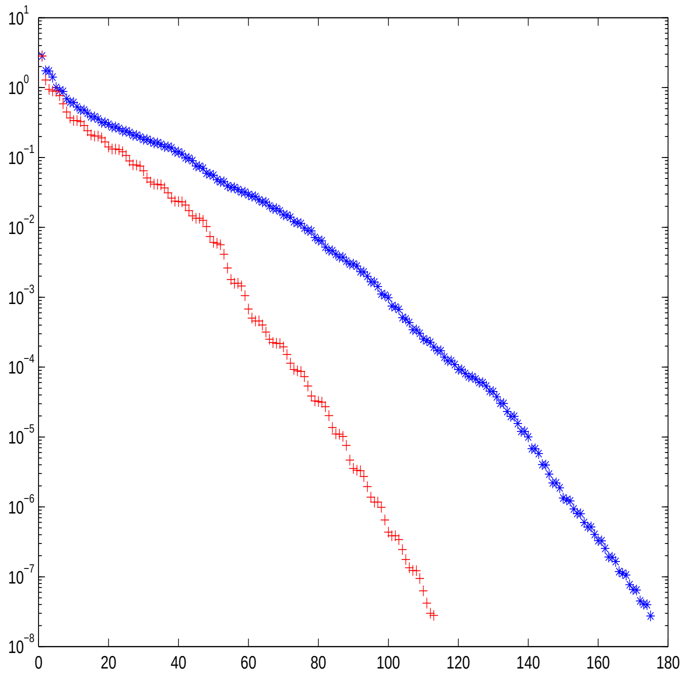
Possible benefits: Allows a tight upper bound for the *condition number*, based solely on norms associated with $\mathcal{A}$:

$$\kappa_2(\mathcal{A}(W)) \leq \kappa_2(\mathcal{A}) + \|W\|_2^2 \|C\|_2^2 + \alpha \|W\|_2,$$

where $\alpha > 0$ depends on $\|\mathcal{A}\|_2, \|\mathcal{A}^{-1}\|_2$ and $\|C\|_2$.

Convergence for $W = \gamma I$

Apply nonpreconditioned GMRES (no restart) to $\mathcal{A}(W)$ (system size is 1856×1856, (1,1) block size is 1344×1344)

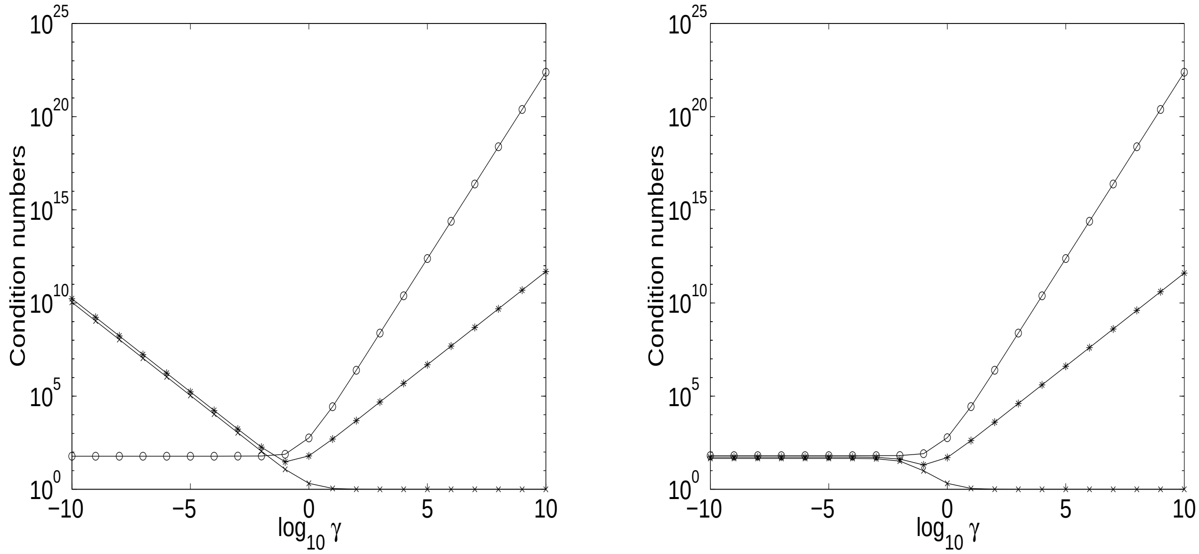


Conv. history for $\gamma = 1$ (red) and $\gamma = 0$ (blue).

The larger γ is, the *more* ill-conditioned the matrix is, but convergence may be faster.

Choice of the parameter

Randomly generated 130 × 130 matrices, with 100 × 100 (1,1) blocks.



(a) semidefinite $A^T A$ (b) positive definite $A^T A$

Condition numbers of the (1,1) block ('*'), the whole KKT matrix ('o'), and the Schur complement ('x'), as a function of γ.

- Sensible choice is crucial if the (1,1) block is singular; the modification of the linear system makes a big difference.

Application of the Uzawa algorithm

Arrow, Hurwicz, Uzawa [1958], *Bank, Welfert & Yserentant* [1990], *Elman & G.* [1994], *Bramble, Pasciak & Vassilev* [1997], ...

For $k = 0, 1, \ldots$
solve $A^T A\mathbf{x}_{k+1} = A^T\mathbf{b} - C\boldsymbol{\lambda}_k$
compute $\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \alpha C^T\mathbf{x}_{k+1}$

- Optimal α depends on extreme eigenvalues of the Schur complement $C^T(A^TA)^{-1}C$.
- No need to find the *exact* solution of the 'inner' system.

In our formulation, *two* parameters are involved: A^TA is actually replaced by $A^TA + \gamma CC^T$, and hence both α and γ need to be determined.

- Can show: any choice of $\gamma > 0$ with $0 < \alpha < 2\gamma$ converges. The choice $\alpha = \gamma$ works well.

2.3 Weighting method

Instead of solving

$$\|\mathbf{b} - A\mathbf{x}\|_2 = \min$$
$$\text{s.t. } C^T\mathbf{x} = \mathbf{0}$$

consider the unconstrained problem:

$$\left(\|\mathbf{b} - A\mathbf{x}\|_2^2 + \mu^2\|C^T\mathbf{x}\|_2^2\right) = \min$$

or

$$\left\| \begin{pmatrix} \mathbf{b} \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} A \\ \mu C^T \end{pmatrix} \mathbf{x} \right\|_2^2 = \min .$$

- Note: For large μ the solution $\hat{\mathbf{x}}(\mu)$ of the unconstrained problem should be a good approximation to the solution $\hat{\mathbf{x}}$ of the constrained problem.

Generalized Singular Value Decomposition (GSVD)

$$\begin{aligned} U^T A X &= \operatorname{diag}(\alpha_1, \ldots, \alpha_m) \\ V^T C^T X &= \operatorname{diag}(\gamma_1, \ldots, \gamma_p) \end{aligned}$$

$U = [\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_m}]$, $V = [\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_p}]$,
$X = [\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}]$.

Solution of the constrained problem

$$\hat{\mathbf{x}} = \sum_{i=p+1}^{n} \frac{\mathbf{u}_i^T \mathbf{b}}{\alpha_i} \mathbf{x}_i.$$

Solution of the unconstrained problem

$$\hat{\mathbf{x}}(\mu) = \sum_{i=1}^{p} \frac{\alpha_i \mathbf{u}_i^T \mathbf{b}}{\alpha_i^2 + \mu^2 \gamma_i^2} \mathbf{x}_i + \hat{\mathbf{x}}.$$

Consequently

$$\hat{\mathbf{x}}(\mu) - \hat{\mathbf{x}} = \sum_{i=1}^{p} \frac{\alpha_i \mathbf{u}_i^T \mathbf{b}}{\alpha_i^2 + \mu^2 \gamma_i^2} \mathbf{x}_i \to 0 \text{ as } \mu^2 \to \infty.$$

(Van Loan)

2.4 Direct method for

$$\|\mathbf{b} - A\mathbf{x}\|_2 = \min, \text{ s.t. } C^T\mathbf{x} = \mathbf{0}.$$

Compute the QR factorization of C

$$Q^T C = \begin{pmatrix} R \\ 0 \end{pmatrix} \begin{matrix} p \\ n-p \end{matrix}$$

and set

$$AQ^T = (\underbrace{A_1}_{p}, \underbrace{A_2}_{n-p}), \quad Q\mathbf{x} = \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} \begin{matrix} p \\ n-p \end{matrix} \quad .$$

Then, the constrained problem becomes

$$\|\mathbf{b} - A_1\mathbf{y} - A_2\mathbf{z}\|_2 = \min, \text{ s.t. } R^T\mathbf{y} = \mathbf{0}.$$

So, $\mathbf{y} = \mathbf{0}$. Let $\hat{\mathbf{z}}$ denote the solution of

$$\|\mathbf{b} - A_2\mathbf{z}\|_2 = \min,$$

then

$$\hat{\mathbf{x}} = Q^T \begin{pmatrix} \mathbf{0} \\ \hat{\mathbf{z}} \end{pmatrix}.$$

3. Least squares with quadratic constraint

Consider the problem of finding $\hat{\mathbf{x}}$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\|_2 = \min$$

subject to the quadratic constraint

$$\|\hat{\mathbf{x}}\|_2^2 = \alpha^2.$$

3.1 Lagrange multipliers

$$\psi(\mathbf{x}, \mu) = \|\mathbf{b} - A\mathbf{x}\|_2^2 + \mu(\|\mathbf{x}\|_2^2 - \alpha^2)$$

grad $\psi = 0$ when

$$\begin{aligned} (A^T A + \mu I)\mathbf{x} &= A^T \mathbf{b} \\ \mathbf{x}^T \mathbf{x} &= \alpha^2 \end{aligned}$$

which leads to the *secular equation*

$$\mathbf{b}^T A(A^T A + \mu I)^{-2} A^T \mathbf{b} - \alpha^2 = 0.$$

SVD/Newton approach

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

$$\sum_{i=1}^{n} \beta_i^2 \frac{\sigma_i^2}{(\sigma_i^2 + \mu)^2} - \alpha^2 = 0$$

Care must be taken to solve this equation. Newton's method can be very delicate.

Quadratic eigenvalue approach

Consider

$$\begin{pmatrix} (A^TA + \mu I)^2 & A^T\mathbf{b} \\ \mathbf{b}^TA & \alpha^2 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \xi \end{pmatrix} = 0.$$

Note,

$$\left((A^TA + \mu I)^2 - \frac{1}{\alpha^2}(\mathbf{b}^TA)^T\mathbf{b}^TA\right)\mathbf{u} = 0.$$

Thus, μ can be found as an eigenvalue of a quadratic eigenvalue problem with $\hat{\mathbf{x}} = \mathbf{u}/\xi$.

Approximating the secular equation

Wanted: solution of

$$f(\mu) \equiv \mathbf{b}^T A(A^T A + \mu I)^{-2} A^T \mathbf{b} = \alpha^2.$$

Note, that

$$\begin{aligned} f(\mu) &= \sum_{i=1}^{n} \beta_i^2 \frac{\sigma_i^2}{(\sigma_i^2 + \mu)^2} \\ &= \int_{\sigma_1}^{\sigma_n} \frac{\sigma^2}{(\sigma^2 + \mu)^2} d\beta(\sigma) \\ &= I(\mu). \end{aligned}$$

This integral may be efficiently bounded by employing the *Lanczos scheme* and exploiting the connection of *modified moments* and the *Gauss - Radau rule*

$$L_k(\mu) \le I(\mu) \le U_k(\mu).$$

Total least squares (TLS)

Consider the problem of finding $\hat{E}$ and $\hat{\mathbf{r}}$ such that

$$\left(\|\hat{E}\|_F^2 + \|\hat{\mathbf{r}}\|_2^2\right) = \min$$

subject to the constraint

$$(A + \hat{E})\hat{\mathbf{x}} = \mathbf{b} + \hat{\mathbf{r}},$$

where the $m \times n$ matrix A and the vector $\mathbf{b}$ are known.

The constraint may be rewritten as follows

$$(A, \mathbf{b}) \begin{pmatrix} \mathbf{x} \\ -1 \end{pmatrix} + (E, \mathbf{r}) \begin{pmatrix} \mathbf{x} \\ -1 \end{pmatrix} = \mathbf{0}$$

or in compact notation

$$(C + F)\mathbf{z} = \mathbf{0}, \text{ with } z_{n+1} = -1.$$

Let rank $C < n + 1$ and let $C = U\Sigma V^T$ denote the SVD of C. Then

$$\begin{pmatrix} \hat{\mathbf{x}} \\ -1 \end{pmatrix} = \mathbf{z} = -\frac{1}{v_{n+1,n+1}} \mathbf{v_{n+1}}.$$

Partial total least squares

Consider the special "error matrix"

$$E = (\underbrace{0}_{p}, E_2)$$

and let again

$$C = (A, \mathbf{b}),\ \ F = (E, \mathbf{r}).$$

Compute the QR-decomposition

$$Q^T(C + F) = \begin{pmatrix} R_{1,1} & \tilde{C}_{1,2} + E_{1,2} \\ 0 & \tilde{C}_{2,2} + E_{2,2} \end{pmatrix}$$

to obtain

$$\begin{aligned} \min \|F\|_F^2 &= \min \|Q^T F\|_F^2 \\ &= \min \left(\|E_{1,2}\|_F^2 + \|E_{2,2}\|_F^2\right). \end{aligned}$$

Then, the SVD

$$\tilde{C}_{2,2} = U \Sigma V^T$$

yields the wanted solution $\hat{\mathbf{x}}^T = (\hat{\mathbf{s}}^T; \hat{\mathbf{t}}^T)$ with

$$\hat{\mathbf{t}} = -\frac{1}{v_{p+1,p+1}} \mathbf{v}_{p+1},\ \ R_{1,1}\hat{\mathbf{s}} = -\tilde{C}_{1,2}\hat{\mathbf{t}}.$$

Regularized total least squares

Note, that the TLS solution is equivalent to

$$\min \frac{\|\mathbf{b} - A\mathbf{x}\|_2^2}{1 + \|\mathbf{x}\|_2^2} = \min \frac{\|C\mathbf{z}\|_2^2}{\|\mathbf{z}\|_2^2} = \sigma_{\min}(C).$$

For the regularized TLS we consider

$$\min \frac{\|\mathbf{b} - A\mathbf{x}\|_2^2}{1 + \mathbf{x}^T V \mathbf{x}}, \text{ subject to } \mathbf{x}^T V \mathbf{x} = \alpha^2,$$

where V is a given symmetric positive definite matrix. Now, let

$$W = \begin{pmatrix} V & 0 \\ 0 & 1 \end{pmatrix} = F^T F$$

and observe that

$$\min \frac{\|\mathbf{b} - A\mathbf{x}\|_2^2}{1 + \mathbf{x}^T V \mathbf{x}} = \min \frac{\|C\mathbf{z}\|_2^2}{\mathbf{z}^T W \mathbf{z}}$$

with $\|\mathbf{z}\|_2^2 = 1 + \alpha^2, \; z_{n+1} = -1.$

Least squares with linear and quadratic constraints

With

$$\mathbf{y} = F\mathbf{z}, B = F^{-T}C^TCF^{-1}, \mathbf{c} = \mathbf{e}_{n+1}^T F^{-1},$$

$$\gamma^2 = 1 + \alpha^2, \text{ and } \beta = -1$$

we may rewrite our regularized TLS problem in terms of a least squares problem with linear **and** quadratic constraints

$$\min \frac{\mathbf{y^T} B\mathbf{y}}{\mathbf{y}^T\mathbf{y}}, \quad \text{s. t. } \|\mathbf{y}\|_2^2 = \gamma^2, \ \mathbf{c}^T\mathbf{y} = \beta.$$

Lagrange multipiers

$$\psi(\mathbf{y}; \lambda, \mu) = \mathbf{y^T} B\mathbf{y} - \lambda(\mathbf{y^T}\mathbf{y} - \gamma^2) - 2\mu(\mathbf{c}^T\mathbf{y} - \beta).$$

grad $\psi = 0$ when

$$B\mathbf{y} - \lambda\mathbf{y} - \mu\mathbf{c} = \mathbf{0}.$$

Introducing the projection matrix

$$P = I - \frac{\mathbf{c}\mathbf{c}^T}{\mathbf{c}^T\mathbf{c}} \text{ and } \mathbf{d} = \frac{\beta\mathbf{c}}{\mathbf{c}^T\mathbf{c}}$$

we arrive at

$$\begin{aligned} (PB - \lambda I)\mathbf{y} &= -\lambda\mathbf{d} \\ \mathbf{y}^T\mathbf{y} &= \gamma^2, \end{aligned}$$

which leads to the secular equation

$$\lambda^2\mathbf{d}^T(PB - \lambda I)^{-T}(PB - \lambda I)\mathbf{d} = \gamma^2.$$

Instead, consider

$$\begin{pmatrix} (PB - \lambda I)(PB - \lambda I)^T & \lambda\mathbf{d} \\ \lambda\mathbf{d}^T & \gamma^2 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \xi \end{pmatrix} = 0.$$

Note,

$$\left((PB - \lambda I)(PB - \lambda I)^T - \frac{\lambda^2}{\gamma^2}\mathbf{d}\mathbf{d}^\mathbf{T}\right)\mathbf{u} = 0.$$

Thus, λ can be found as an eigenvalue of a quadratic eigenvalue problem with $\hat{\mathbf{y}} = \mathbf{u}/\xi$.