FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG
INSTITUT FÜR INFORMATIK (MATHEMATISCHE MASCHINEN UND DATENVERARBEITUNG)

Lehrstuhl für Informatik 10 (Systemsimulation)

Iterative Solution of Weighted Least Squares Problems with Applications to Rigid Multibody Dynamics

C. Popa and T. Preclik

Technical Report 10-10

Iterative Solution of Weighted Least Squares Problems with Applications to Rigid Multibody Dynamics

C. Popa*, T. Preclik†

30th November 2010

Abstract.

Riley suggested an iterated regularization scheme back in 1956. It can be used to compute solutions of linear least squares problems of minimum Euclidean norm. In this report this scheme is generalized to solutions of weighted minimum Euclidean norm. Convergence results are obtained with the aid of a generalized singular value decomposition and numerical experiments with applications in rigid multibody dynamics are carried out.

1 Introduction

A timestep problem in rigid multibody dynamics can usually be formulated with the help of the Jacobian J, the mass matrix M and the impulses x.

$$JM^{-1}J^T x + J^T b = 0 \tag{1}$$

This linear system of equations (LSE) is possibly *overconstrained* that is *underdetermined*. Though the mass matrix is symmetric positive-definite (SPD), the Jacobian is typically rectangular and rank-deficient, where J can be fat ($m \leq n$) as well as tall ($m \geq n$). As was shown in [9] the system can be regularized by relaxing the rigidity assumption. The regularization leads to

$$JM^{-1}J^T x + J^T b = -sDx, \tag{2}$$

where D is positive and diagonal. The regularization entails that the constraints are violated. To deal with this problem one can uniformly increase the stiffness at the constraints which corresponds to driving s towards 0 in (2). This poses several questions: How can we characterize the solution in the stiff limit and how can we efficiently compute it? To approach the answers the limit process is analyzed with the aid of singular value decompositions (SVD).

2 The Singular Value Decomposition

It seems that the SVD has a very long history (see [11] and [4], page 74: "Notes and references for sec. 2.5"). In what follows we shall present the version from [4], page 70 (Theorem 2.5.2).

*Ovidius University of Constanta, Faculty of Mathematics and Informatics, cpopa@univ-ovidius.ro

†University of Erlangen-Nürnberg, Computer Science 10, System Simulation, tobias.preclik@informatik.uni-erlangen.de

Theorem 2.1. *(SVD)*
For any $A \in \mathcal{M}_{m\times n}$ *with* $1 \leq rank(A) = r \leq \min\{m,n\}$, *there exist orthogonal matrices* $U : m \times m$ *and* $V : n \times n$ *such that*

$$U^T AV = \Sigma = \text{diag}(\sigma_1, ..., \sigma_r, 0, ..., 0) \quad \text{and} \quad \sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0. \tag{3}$$

Some properties of the SVD are given in the following results (proofs can be found in [4] and [2]). By $\mathcal{N}(E), \mathcal{R}(E)$ we shall denote the null-space, respectively, the range of a matrix E.

Proposition 1. *(i)* $\sigma_1^2, \ldots, \sigma_r^2$ *are the nonzero eigenvalues of the matrices* $A^T A$ *and* AA^T.
(ii) An SVD decomposition for A^T *is given by*

$$V^T A^T U = \Sigma^T. \tag{4}$$

(iii) If $U = \text{col}(u_1, ..., u_m), V = \text{col}(v_1, ..., v_n)$, *then the following hold*

$$Av_i = \sigma_i u_i,\ A^T u_i = \sigma_i v_i, A^T A v_i = \sigma_i^2 v_i,\ AA^T u_i = \sigma_i^2 u_i, i = 1, \ldots, r, \tag{5}$$

$$\mathcal{N}(A) = \text{sp}(v_{r+1}, \ldots, v_n), \quad \mathcal{R}(A) = \text{sp}(u_1, \ldots, u_r) \tag{6}$$

$$\mathcal{N}(A^T) = \text{sp}(u_{r+1}, \ldots, v_m), \quad \mathcal{R}(A^T) = \text{sp}(v_1, \ldots, v_r) \tag{7}$$

and

$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^T. \tag{8}$$

(iv) The $n \times m$ *matrix* A^+ *defined by*

$$A^+ = V\Sigma^+ U^T, \tag{9}$$

where Σ^+ *is the* $n \times m$ *diagonal matrix given by (see (3))*

$$\Sigma^+ = \text{diag}(\frac{1}{\sigma_1}, \ldots, \frac{1}{\sigma_r}, 0, \ldots, 0) \tag{10}$$

is the Moore-Penrose pseudoinverse of A.

Moreover, the SVD gives us complete information about the solution set of the linear least squares problem

$$\| Ax^* - b \| = \min\{\| Ax - b \|, x \in I\!R^n\}, \tag{11}$$

denoted by $LSS(A; b)$ (x_{LS} will be its minimal norm solution).

Proposition 2. *With the notations from Theorem 2.1 and Proposition 1 we have*

$$LSS(A; b) = \left\{ \sum_{i=1}^{r} \frac{\langle u_i, b\rangle}{\sigma_i} v_i + \sum_{i=r+1}^{n} \alpha_i v_i,\ \alpha_i \in I\!R \right\} \tag{12}$$

and

$$x_{LS} = \sum_{i=1}^{r} \frac{\langle u_i, b\rangle}{\sigma_i} v_i. \tag{13}$$

Proof. We present a proof as a model for some specific working techniques that will be used later in this paper. Let $x \in I\!R^n$ be arbitrary fixed. Because the Euclidean norm is invariant under orthogonal transformation, from (3) we get

$$\| Ax - b \|^2 = \| (U^T AV)(V^T x) - U^T b \|^2 = \| \Sigma z - w \|^2 = \sum_{i=1}^{r} (\sigma_i z_i - w_i)^2 + \sum_{i=r+1}^{m} w_i^2, \tag{14}$$

where we denoted

$$z = V^T x \Leftrightarrow x = Vz \quad \text{and} \quad w = U^T b \Leftrightarrow w_i = \langle u_i, b\rangle. \tag{15}$$

The minimal value of the expression in (14) is obtained for $z_i = \frac{w_i}{\sigma_i} = \frac{\langle u_i, b\rangle}{\sigma_i}$. Then (12) results by also using the equality $x = Vz$. If x^* is an arbitrary element from the set in the right hand side of the equality (12) we have

$$\parallel x^* \parallel^2 = \sum_{i=1}^{r} \left(\frac{\langle u_i, b\rangle}{\sigma_i} \right)^2 + \sum_{i=r+1}^{n} (z_i)^2 \geq \sum_{i=1}^{r} \left(\frac{\langle u_i, b\rangle}{\sigma_i} \right)^2 .$$

This gives us (13) and completes the proof. □

Note 1. Formula (13) can be also obtained from (9)-(10) and the equality $x_{LS} = A^+ b$.

3 The generalized singular value decomposition (GSVD)

There were several reasons for constructing the GSVD (see the comments in [2, 4, 6, 7] and references therein); among them we shall mention two in which we are directly interested in this paper: constrained and weighted versions of the problem (11), and Tikhonov-like regularized formulations of it. In this last case, instead of (11) we consider the problem: find $x^* \in I\!R^n$ such that

$$\parallel Ax^* - b \parallel^2 + s^2 \parallel Bx^* \parallel^2 = \min\{\parallel Ax - b \parallel^2 + s^2 \parallel Bx \parallel^2, x \in I\!R^n\}, \tag{16}$$

where B is an $n \times n$ invertible matrix and $s \geq 0$ the regularization parameter. We firstly observe that the problem (16) is equivalent with the following one (which is of the form (11)): find $x^* \in I\!R^n$ such that

$$\parallel \begin{bmatrix} A \\ sB \end{bmatrix} x^* - \begin{bmatrix} b \\ 0 \end{bmatrix} \parallel^2 = \min\{\parallel \begin{bmatrix} A \\ sB \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \parallel^2, x \in I\!R^n\} \tag{17}$$

If $B = I$ we can give complete information (of the form (12)-(13)) about the set of all solutions of the problem (11), by just using the SVD of A and the technique from the proof of Proposition 2. For a general matrix B we can reduce the problem (17) to the case $B = I$, by introducing the substitution $Bx = y$. In this way (17) becomes equivalent (with respect to the relation $x = B^{-1}y$) with the problem: find $y^* \in I\!R^n$ such that

$$\parallel \begin{bmatrix} AB^{-1} \\ sI \end{bmatrix} y^* - \begin{bmatrix} b \\ 0 \end{bmatrix} \parallel^2 = \min\{\parallel \begin{bmatrix} AB^{-1} \\ sI \end{bmatrix} y - \begin{bmatrix} b \\ 0 \end{bmatrix} \parallel^2, y \in I\!R^n\}. \tag{18}$$

But, for (18) we would need an SVD for the matrix AB^{-1}, which may be difficult and not indicated to construct (e.g. for ill conditioned matrices B). The GSVD can provide us information similar to those from Theorem 2.1, for the matrix AB^{-1}, but without constructing it explicitly. In what follows we shall present the GSVD in the particular case in which **the matrix B is square ($n \times n$) and invertible** (because only this case will interest us in the rest of the paper).
Note 2. The first version of the GSVD was given in [6], in the particular case $m \geq n$, whereas its general version, for any m and n can be found in [7].

Case 1: $m \geq n$. In this case we shall use the statement and notations from Theorem 4.2.2, page 157 in [2]. Because the matrix B is square and nonsingular we have $p = n = k = q$ and there exist the orthogonal matrices $U_A : m \times m$, $U_B : n \times n$ and the nonsingular one $Z : n \times n$ such that

$$U_A^T A = D_A Z, \quad U_B^T B = D_B Z, \tag{19}$$

where $D_A : m \times n, D_B : n \times n$ are given by

$$D_A = \begin{bmatrix} \alpha_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_n \\ 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix}, \quad D_B = \begin{bmatrix} \beta_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \beta_n \end{bmatrix}, \tag{20}$$

and

$$0 \leq \alpha_1 \leq \cdots \leq \alpha_n < 1, \ 1 \geq \beta_1 \geq \cdots \geq \beta_n > 0, \ \alpha_i^2 + \beta_i^2 = 1, \forall i = 1, \dots, n. \tag{21}$$

In what follows we will be concerned with the answer to the following problem: **How does** $r = rank(A) \leq n$ **appear in the above decomposition (19)-(21)?** For this, we go into the proof of Theorem 4.2.2 page 157 in [2] and observe that it starts with an SVD of the extended matrix $\begin{bmatrix} A \\ B \end{bmatrix}$ of the form

$$Q^T \begin{bmatrix} A \\ B \end{bmatrix} P = \begin{bmatrix} \Sigma \\ 0 \end{bmatrix}, \tag{22}$$

where Q and P are orthogonal matrices of order $(m+n)$ and n, respectively and Σ a diagonal one of order n given by (because $rank(\begin{bmatrix} A \\ B \end{bmatrix}) = n$)

$$\Sigma = diag(\sigma_1, \dots, \sigma_n), \quad \sigma_1 \geq \cdots \geq \sigma_n > 0. \tag{23}$$

From (22) we then get

$$\begin{bmatrix} A \\ B \end{bmatrix} P = Q \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} = \begin{bmatrix} Q_{11}\Sigma \\ Q_{21}\Sigma \end{bmatrix}, \tag{24}$$

where Q_{11}, Q_{21} have dimensions $m \times n$, respectively $n \times n$ and Q_{21} is invertible. From (24) we obtain $AP = Q_{11}\Sigma$, thus (because P and Σ are invertible; see **Note 3** below)

$$\mathcal{R}(A) = \mathcal{R}(AP) = \mathcal{R}(Q_{11}\Sigma) = \mathcal{R}(Q_{11}) \ \Rightarrow \ rank(Q_{11}) = rank(A) = r. \tag{25}$$

Note 3. The value of $rank(E)$ for a matrix E is the dimension of $\mathcal{R}(E)$. If $E : m \times n$ is arbitrary and $F : n \times n$ is invertible we have $\mathcal{R}(E) = \mathcal{R}(EF)$, thus $rank(E) = rank(EF)$.

We now go back again at the proof of Theorem 4.2.2 page 157 in [2], and observe that the numbers $\alpha_i, \beta_i, i = 1, \dots, n$ and the matrices U_A, U_B from (19)-(21) are given by a CS-decomposition (Theorem 4.2.1, page 155) for $\begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix}$ of the form

$$Q_{11} = U_A \begin{bmatrix} C \\ 0 \end{bmatrix} V^T, \quad Q_{21} = U_B S V^T \tag{26}$$

($V : n \times n$ orthogonal), where the $n \times n$ diagonal matrices C, S are given by

$$C = diag(\alpha_1, \dots, \alpha_n), \quad S = diag(\beta_1, \dots, \beta_n), \tag{27}$$

with α_i, β_i as in (21) and $D_A = \begin{bmatrix} C \\ 0 \end{bmatrix}, D_B = S$. Then, from (25)-(27) it results that $r = rank(A) = rank(Q_{11}) = rank(diag(\alpha_1, \dots, \alpha_n))$ which means that only r values α_i are strictly positive, i.e. according to the inequalities in (21)

$$0 = \alpha_1 = \cdots = \alpha_{n-r} < \alpha_{n-r+1} \le \cdots \le \alpha_n \le 1 \tag{28}$$

and

$$1 = \beta_1 = \cdots = \beta_{n-r} > \beta_{n-r+1} \ge \cdots \ge \beta_n > 0. \tag{29}$$

Moreover, from (19) we obtain

$$AB^{-1} = U_A D_A Z Z^{-1} D_B^{-1} U_B^T = U_A D_A D_B^{-1} U_B^T, \tag{30}$$

which is a singular value decomposition of the matrix AB^{-1}. From here and (28)-(29) it results that the ratios

$$\frac{\alpha_i}{\beta_i} > 0, i = n-r+1, \ldots, n, \tag{31}$$

represent the r nonzero singular values of the matrix AB^{-1} ($rank(AB^{-1}) = rank(A) = r$).

Note 4. The inequalities in (21) can be reversed according to the proof of the CS-decomposition Theorem 4.2.1 from [2], page 155 (by reordering in an appropriate way the angles θ_i, see page 156). In this case we would have in (21)

$$1 \ge \alpha_1 \ge \cdots \ge \alpha_n \ge 0,\ 0 < \beta_1 \le \cdots \le \beta_n \le 1,\ \alpha_i^2 + \beta_i^2 = 1, \forall i = 1, \ldots, n, \tag{32}$$

in (28)-(29)

$$1 \ge \alpha_1 \ge \cdots \ge \alpha_r > \alpha_{r+1} = \cdots = \alpha_n = 0, \tag{33}$$

$$0 < \beta_1 \le \cdots \le \beta_r < \beta r + 1 = \cdots = \beta_n = 1, \tag{34}$$

and in (31)

$$\frac{\alpha_i}{\beta_i} > 0,\ i = 1, \ldots, r. \tag{35}$$

Case 2: $m \le n$. In this case we will use the statement from the Theorem on page 399 in [7] (see also [1]).

Because the matrix B is square and nonsingular we there have $p = n = k$ and there exist the orthogonal matrices $U_A : m \times m$, $U_B : n \times n, Q : n \times n$ and the nonsingular one $Z : n \times n$ such that

$$U_A^T A Q = D_A Z, \quad U_B^T B Q = D_B Z, \tag{36}$$

with $D_A : m \times n, D_B : n \times n$ are of the form

$$D_A = \begin{bmatrix} \alpha_1 & & & 0 & \ldots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \alpha_m & 0 & \ldots & 0 \end{bmatrix}, \quad D_B = \begin{bmatrix} \beta_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \beta_n \end{bmatrix}, \tag{37}$$

and

$$1 > \alpha_1 \ge \cdots \ge \alpha_m \ge 0,\ 0 < \beta_1 \le \cdots \le \beta_n \le 1,\ \alpha_i^2 + \beta_i^2 = 1, \forall i = 1, \ldots, m. \tag{38}$$

In what follows we will be concerned with the answer to a similar problem as in **Case 1**: **How does $r = rank(A) \le n$ appear in the above decomposition (36)-(38)?** In this respect, firstly let us note that, as in **Note 3** we obtain, according to (36)

$$rank(D_A) = rank(D_A Z) = rank(U_A^T A Q). \tag{39}$$

Note 5. Let $E : m \times n$ be arbitrary and $F : n \times n$ invertible. As in **Note 3** we obtain $rank(E^T F^T) = rank(E^T)$. But, for any matrix G, $rank(G) = rank(G^T)$, thus $rank(FE)$=$rank(E^T F^T)$=$rank(E^T)$=$rank(E)$.

From the above **Note 5** and (39) we get

$$rank(D_A) = rank(D_A Z) = rank(U_A^T (AQ)) = rank(AQ) = rank(A) = r. \tag{40}$$

From (40) and (37)-(38) it then holds

$$1 > \alpha_1 \geq \cdots \geq \alpha_r > \alpha_{r+1} = \cdots = \alpha_m = 0, \tag{41}$$

and we obtain also in this case a similar result with the one in **Note 4**.

Conclusion. From the above considerations, for any m and n we get the following form of the GSVD of the pair (A, B), $rank(A) = r \leq \min(m, n)$ in the particular case $B : n \times n$ invertible

$$U_A^T A X = D_A = diag(\alpha_1, \ldots, \alpha_r, 0, \ldots, 0),$$

$$U_B^T B X = D_B = diag(\beta_1, \ldots, \beta_n), \tag{42}$$

with $U_A : m \times m$ and $U_B : n \times n$ orthogonal, $X : n \times n$ invertible, $D_A : m \times n$, $D_B : n \times n$,

$$1 > \alpha_1 \geq \cdots \geq \alpha_r > 0, \quad 0 < \beta_1 \leq \cdots \leq \beta_r < \beta_{r+1} = \cdots = \beta_n = 1,$$

$$\alpha_i^2 + \beta_i^2 = 1, i = 1, \ldots, r, \tag{43}$$

and the ratios in (35) are the nonzero singular values of the matrix AB^{-1}.

4 Weighted Least Squares Problems from Rigid Multibody Dynamics

We consider an arbitrary $m \times n$ matrix A with $rank(A) = r \leq \min(m, n)$, $b \in I\!R^m$ and the linear least squares problem (11) arising from a time-step problem in rigid multibody dynamics. It is well known that (11) is equivalent with the associated normal equation: find $x^* \in I\!R^n$ such that

$$A^T A x^* = A^T b. \tag{44}$$

Such a system is possibly rank-deficient which is fundamentally connected to the rigidity assumption. When relaxing the rigidity assumption, the multibody system can be described by the following regularized variant of (44), where the stiffness is controlled by the term sD:

$$A^T A x^*(s) + A^T b = -sDx^*(s), \tag{45}$$

where D is an SPD matrix and $s \in I\!R, s > 0$. When s tends to 0, the solution approaches the completely rigid system behaviour. Though a diagonal D with positive entries would suffice to describe regularized multibody systems we will investigate the more general case. The problem (45) has a unique solution, $x^*(s)$, because the system matrix $A^T A + sD$ is SPD (indeed we have $\langle (A^T A + sD)x, x \rangle = \parallel Ax \parallel^2 + s\langle Dx, x \rangle > 0, \forall x \neq 0$). Let $J : I\!R^n \longrightarrow I\!R$ be defined by

$$J(x) = \langle A^T A x, x \rangle - 2\langle A^T b, x \rangle + s\langle Dx, x \rangle. \tag{46}$$

Its gradient is given by

$$J'(x) = 2A^T A x - 2A^T b + 2sDx, \tag{47}$$

which means that the problem (45) is equivalent with

$$J(x^*(s)) = \min_{x \in I\!R^n} J(x). \tag{48}$$

We shall now consider another functional $G : I\!R^n \longrightarrow I\!R$, defined by

$$G(x) = J(x) + \parallel b \parallel^2 . \tag{49}$$

From (46) and (49) we get

$$G(x) = \| Ax - b \|^2 + (\sqrt{s})^2 \| D^{\frac{1}{2}} x \|^2 . \qquad (50)$$

Moreover, from (49) (because the additional term $\| b \|^2$ is constant) we obtain that the problem (48) is equivalent with

$$G(x^*(s)) = \min_{x \in I\!R^n} G(x). \qquad (51)$$

Thus, solving (45) is equivalent with solving (51). Then, we firstly observe that the problem (51) is a Tikhonov - like regularization, and it can be written as (see also (16)-(17))

$$G(x^*(s)) = \min_{x \in I\!R^n} \| \begin{bmatrix} A \\ \sqrt{s} D^{\frac{1}{2}} \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \|^2 . \qquad (52)$$

Let $x^*(s)$ be its unique solution. We will use a Generalized Singular Value Decomposition of the matrix $\begin{bmatrix} A \\ D^{\frac{1}{2}} \end{bmatrix}$ as in (42)-(43) (for $B = D^{\frac{1}{2}}$)

$$U_A^T A X = diag(\alpha_1, \dots, \alpha_r, 0, \dots, 0) = D_A,$$
$$U_B^T D^{\frac{1}{2}} X = diag(\beta_1, \dots, \beta_n) = D_B. \qquad (53)$$

Then, by introducing (53) in (52) and using the orthogonality of U and V and the invertibility of X we successively obtain (see e.g. the proof of Proposition 2)

$$\min_{x \in I\!R^n} \| \begin{bmatrix} A \\ \sqrt{s} D^{\frac{1}{2}} \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \|^2 =$$

$$\min_{x \in I\!R^n} \| \begin{bmatrix} U_A^T & 0 \\ 0 & U_B^T \end{bmatrix} \left(\begin{bmatrix} A \\ \sqrt{s} D^{\frac{1}{2}} \end{bmatrix} X (X^{-1} x) - \begin{bmatrix} b \\ 0 \end{bmatrix} \right) \|^2 =$$

$$\min_{z \in I\!R^n} \| \begin{bmatrix} D_A \\ \sqrt{s} D_B \end{bmatrix} z - \begin{bmatrix} w \\ 0 \end{bmatrix} \|^2 =$$

$$\sum_{i=1}^{r} (\alpha_i z_i - w_i)^2 + \sum_{i=r+1}^{m} w_i^2 + \sum_{i=1}^{n} s \beta_i^2 z_i^2 = E(z), \qquad (54)$$

where $z = (z_1, \dots, z_n)^T = X^{-1} x$, $w = (w_1, \dots, w_n)^T = U_A^T b$. The values of z_i which ensures the minimal value of the function $E(z)$ are obtained by solving the system $\frac{\partial E}{\partial z_i}(z) = 0, i = 1, \dots, n$. These are

$$z_i = \frac{w_i \alpha_i}{\alpha_i^2 + s \beta_i^2}, i = 1, \dots, r, \quad z_i = 0, i = r+1, \dots, n. \qquad (55)$$

Thus, the unique solution $x^*(s)$ of the problem (51) will be given by

$$x^*(s) = Xz = \sum_{i=1}^{n} z_i X^i = \sum_{i=1}^{r} \frac{w_i \alpha_i}{\alpha_i^2 + s \beta_i^2} X^i, \qquad (56)$$

where $X^i \in I\!R^n$ is the i-th column of the matrix X. Now let us consider the weighted LSS problem

$$\| D^{\frac{1}{2}} x_D^* \|^2 = \min_{x \in I\!R^n, s.t. \| Ax - b \| = \min!} \| D^{\frac{1}{2}} x \|^2, \qquad (57)$$

with the unique solution $x_D^* \in I\!R^n$. That is, we must find $x_D^* \in LSS(A; b)$ for which $\| D^{\frac{1}{2}} x_D^* \|$ is minimal. By using again the GSVD (53) we obtain

$$\| Ax - b \|^2 = \sum_{i=1}^{r} (\alpha_i z_i - w_i)^2 + \sum_{i=r+1}^{n} w_i^2$$

from which we get

$$LSS(A;b) = \{x^* = Xz^*, z_i^* = \frac{w_i}{\alpha_i}, i = 1, \ldots, r; z_i^* \in I\!R, i = r+1, \ldots, n\}. \tag{58}$$

Now, $U_B^T D^{\frac{1}{2}} X = D_B$, i.e. $D^{\frac{1}{2}} = U_B D_B X^{-1}$, thus, for an $x^* \in LSS(A;b)$ we obtain

$$\| D^{\frac{1}{2}} x^* \|^2 = \| D_B z^* \|^2 = \sum_{i=1}^{r} (\frac{\beta_i w_i}{\alpha_i})^2 + \sum_{i=r+1}^{n} (\beta_i z_i^*)^2, \tag{59}$$

which has its minimal value for $z_i^* = 0, i = r+1, \ldots, n$ and gives us

$$x_D^* = Xz^* = \sum_{i=1}^{r} \frac{w_i}{\alpha_i} X^i. \tag{60}$$

Then, from (56) and (60) the desired equality results, i.e.

$$\lim_{s \to 0} x(s) = x_D^*. \tag{61}$$

Remark 1. *The "stiffness control" of the rigid multibody system, sD, proposed in (45) tells us that the system approaches the completely rigid system behaviour in an uniform way, with respect to the parameter s. A "more realistic" possibility would be to relax this idea by considering, e.g. a diagonal positive definite matrix $D = diag(d_1, \ldots, d_n), d_i > 0$ such that, for some matrix norm $\| \cdot \|$ we have*

$$\| D \| \to 0. \tag{62}$$

The question would then be: what happens with the previous results involving the solutions $x^*(s)$ and x_D^* ? Firstly we have to observe that the parameter s does not any more exist and the functional G in the problem (51) will become (see also (50))

$$G(x) = \| Ax - b \|^2 + \| D^{\frac{1}{2}} x \|^2 . \tag{63}$$

Then, the computations from (52) - (55) can be easily adapted to this new case and we get as the unique solution for the problem (51) with G from (63) the vector (see (56) for comparison)

$$x^*(D) = \sum_{i=1}^{r} \frac{w_i \alpha_i}{\alpha_i^2 + \beta_i^2} X^i. \tag{64}$$

Now, for an (arbitrary fixed) matrix D as before the computations in (57)-(59) rest valid and we get the unique solution x_D^* of the prooblem (57) of the form (60). But, now comes the "tricky" thing: the equality (61) can not be written any more ! This is because for $\| D \| \to 0$ the problem (57) will no more exist ! This can be also seen from the GSVD relation (53), $U_B^T D^{\frac{1}{2}} X = diag(\beta_1, \ldots, \beta_n) = D_B$; indeed, if in (62) we consider the spectral or Frobenius norm (which are invariant under the orthogonal transformation U_B) we get

$$0 = \lim \| D \| = \lim \| D^{\frac{1}{2}} \| = \lim \| U_B D_B X^{-1} \| = \lim \| D_B X^{-1} \|, \tag{65}$$

i.e. the matrix product $D_B X^{-1}$ tends to 0, but not separately D_B and/or X^{-1} !
Thus, unfortunately, using the "more realistic" approach proposed before will drive us to an "bottleneck-like" situation, in which we have no more control on the solutions $x^*(D)$ and x_D^*.

5 Iteratively Solving Weighted Least Squares Problems

In this section we analyse an iterative method for the efficient numerical solution of the normal equation (44). For D and s as in Section 3 we can write it in the form

$$x^0 \in I\!R^n, (A^T A + sD) x^{k+1} = sDx^k + A^T b, \quad k \geq 0. \tag{66}$$

Note 6. The convergence of the iteration (66) was analysed in [10] in the case $D = I$ and A^TA invertible, and in [3] for $D = I$, but a general $m \times n$ matrix A. In what follows we shall analyse the convergence of the iteration (66) for a general $m \times n$ matrix A and an SPD matrix D. In our considerations we shall use the proof ideas in [3], but with respect to a GSVD of the pair $(A, D^{\frac{1}{2}})$, where $D^{\frac{1}{2}}$ is the square root of D. Firstly we shall observe that, if we write (66) as

$$x^0 \in I\!R^n, \quad x^{k+1} = Gx^k + c \tag{67}$$

where

$$G = s(A^TA + sD)^{-1}D, \quad c = (A^TA + sD)^{-1}A^Tb \tag{68}$$

then, for $x^0 = 0$ we obtain

$$x^k = (G^{k-1} + G^{k-2} + \cdots + I)c \tag{69}$$

According to the GSVD decomposition (53) it results

$$G = sX(D_A^TD_A + sD_B^2)^{-1}D_B^2X^{-1}, c = X(D_A^TD_A + sD_B^2)^{-1}D_A^Tw, w = U_A^Tb, \tag{70}$$

thus

$$G^j = s^jX(D_A^TD_A + sD_B^2)^{-j}(D_B^2)^jX^{-1}, j = 0, 1, \dots, k-1 \tag{71}$$

The $n \times n$ matrices $D_A^TD_A$ and D_B^2 are diagonal, of the form (see (53))

$$D_A^TD_A = \begin{pmatrix} \alpha_1^2 & & & & & \cdots & 0 \\ & \ddots & & & & & \vdots \\ & & \alpha_r^2 & & & & \\ & & & 0 & & & \\ \vdots & & & & \ddots & & \\ 0 & \cdots & & & & & 0 \end{pmatrix}, D_B^2 = \begin{pmatrix} \beta_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \beta_n^2 \end{pmatrix}. \tag{72}$$

From (70)-(72) we then obtain

$$s^j(D_A^TD_A + sD_B^2)^{-j}(D_B^2)^j = \begin{pmatrix} \left(\frac{s\beta_1^2}{\alpha_1^2+s\beta_1^2}\right)^j & & & & \cdots & 0 \\ & \ddots & & & & \vdots \\ & & \left(\frac{s\beta_r^2}{\alpha_r^2+s\beta_r^2}\right)^j & & & \\ & & & 1 & & \\ \vdots & & & & \ddots & \\ 0 & \cdots & & & & 1 \end{pmatrix} \tag{73}$$

and

$$(D_A^TD_A + sD_B^2)^{-1}D_A^Tw = \begin{pmatrix} \frac{w_1\alpha_1}{\alpha_1^2+s\beta_1^2} \\ \vdots \\ \frac{w_r\alpha_r}{\alpha_r^2+s\beta_r^2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{74}$$

We observe that

$$0 < \frac{s\beta_i^2}{\alpha_i^2 + s\beta_i^2} < 1 \quad \forall i = 1, \dots, r. \tag{75}$$

By introducing all these formulas in (69) we get (by also using (75))

$$x^k = X \begin{pmatrix} \sum_{j=0}^{k-1} \left(\frac{s\beta_1^2}{\alpha_1^2+s\beta_1^2}\right)^j \cdot \frac{w_1\alpha_1}{\alpha_1^2+s\beta_1^2} \\ \vdots \\ \sum_{j=0}^{k-1} \left(\frac{s\beta_r^2}{\alpha_r^2+s\beta_r^2}\right)^j \cdot \frac{w_r\alpha_r}{\alpha_r^2+s\beta_r^2} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = X \begin{pmatrix} \left(1 - \left(\frac{s\beta_1^2}{\alpha_1^2+s\beta_1^2}\right)^k\right) \frac{w_1}{\alpha_1} \\ \vdots \\ \left(1 - \left(\frac{s\beta_r^2}{\alpha_r^2+s\beta_r^2}\right)^k\right) \frac{w_r}{\alpha_r} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{76}$$

$$= \sum_{i=1}^{r} \left(1 - \left(\frac{s\beta_i^2}{\alpha_i^2 + s\beta_i^2}\right)^k\right) \frac{w_i}{\alpha_i} \cdot X^i.$$

Using again (75), (76) and (60) we obtain

$$\lim_{k\to\infty} x^k = x_D^*, \tag{77}$$

i.e. the sequence $\{x^k\}_{k\geq 0}$ generated with the iteration (66) with $x^0 = 0$ and $s > 0$ fixed arbitrary, converges to the unique solution of the weighted least squares problem (57). Moreover, from (60) and (76) it results $x_D^*, x^k \in \text{sp}\{x^1, \ldots, x^r\}$, thus for the error vector $e^k = x^k - x_D^*$ we have

$$e^k = \sum_{i=1}^{r} e_i^k X^i \in \text{sp}\{X^1, \ldots, X^r\}, \quad e_i^k \in I\!R. \tag{78}$$

If we define $f^k = X^{-1}e^k$, from (78) it results

$$f^k = \sum_{i=1}^{r} e_i^k X^{-1} X^i = (e_1^k, \ldots, e_r^k, 0, \ldots, 0)^T \in I\!R^n. \tag{79}$$

Now, because $e^k = Ge^{k-1}$ (see (67) and the relation $x_D^* = Gx_D^* + c$), by using (70), (73) (for $j = 1$) and (79) we get

$$f^k = X^{-1}e^k = X^{-1}Ge^{k-1} = sX^{-1}X(D_A^T D_A + sD_B^2)^{-1} D_B^2 X^{-1} e^{k-1}$$

$$= \begin{pmatrix} \frac{s\beta_1^2}{\alpha_1^2+s\beta_1^2} & & & & \cdots & 0 \\ & \ddots & & & & \vdots \\ & & \frac{s\beta_r^2}{\alpha_r^2+s\beta_r^2} & & & \\ & & & 1 & & \\ \vdots & & & & \ddots & \\ 0 & \cdots & & & & 1 \end{pmatrix} \begin{pmatrix} e_1^{k-1} \\ \vdots \\ e_r^{k-1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{80}$$

By taking the Euclidean norm of (80) we obtain

$$\|f^k\|^2 = \sum_{i=1}^{r} \left(\frac{s\beta_i^2}{\alpha_i^2 + s\beta_i^2}\right)^2 \left(e_i^{k-1}\right)^2 \leq \max_{1\leq i\leq r} \left(\frac{s\beta_i^2}{\alpha_i^2 + s\beta_i^2}\right)^2 \cdot \|f^{k-1}\|^2$$
$$= \left(\max_{1\leq i\leq r} \frac{s}{s + \left(\frac{\alpha_i}{\beta_i}\right)^2}\right)^2 \|f^{k-1}\|^2 \leq \frac{s^2}{\left(s + \min_{1\leq i\leq r} \left(\frac{\alpha_i}{\beta_i}\right)^2\right)^2} \|f^{k-1}\|^2, \tag{81}$$

which gives us information about the step reduction factor (with respect to the Euclidean norm) of the weighted errors f^k from (79). If we consider the relations (43) we obtain that

$$\min_{1\leq i\leq r} \left(\frac{\alpha_i}{\beta_i}\right)^2 = \frac{\alpha_r^2}{\beta_r^2} = \frac{1-\beta_r^2}{\beta_r^2} = \frac{\alpha_r^2}{1-\alpha_r^2} =: \mu \tag{82}$$

Note 7. From (81)-(82) we obtain

$$\|f^k\| \leq \frac{s}{s+\mu}\|f^{k-1}\| \tag{83}$$

Thus, if $s \approx \mu$ we obtain a step error reduction factor $\approx \frac{1}{2}$. If $s \gg \mu$, we have $\frac{s}{s+\mu} \approx 1$ which is not a good error reduction factor per iteration. If $s \ll \mu$ it will determine the increase of the condition number of the matrix $A^TA + sD$ in (66), which makes the computation of x^{k+1} difficult.

6 Numerical experiments

In the following numerical experiments are performed with the iteration scheme described in section 5. The test problems are time step problems arising in rigid multibody dynamics. The test scenarios are described in detail in [8]. The test cases contain only bilateral constraints so that the resulting systems are LSEs of the following kind:

$$A^TAx = A^Tb \tag{84}$$

For each test case the iteration was executed for an unweighted regularization where $D = I$ and for a weighted regularization where D was chosen randomly. The weights were computed by generating a standard normally-distributed random variable, taking the absolute value and adding 1, such that all weights were greater equal and close to 1. The weighted and the unweighted tests were executed with three different values for the parameter s. The parameter was chosen such that an error reduction at least of a factor $f \in \{0.1, 0.5, 0.9\}$ was achieved. From (83) we can conclude that $s = \frac{f}{1-f}\mu$, where μ is the smallest nonzero singular value of AD^{-1} squared.

6.1 Test Case: Well

In the well test case $A \in I\!R^{1200\times 6240}$ has full row-rank ($\operatorname{rank} A = 1200$) and thus $A^TA \in I\!R^{6240\times 6240}$ is rank deficient ($\operatorname{rank} A^TA = \operatorname{rank} A = 1200$). A singular value decomposition of A reveals that the smallest nonzero singular value σ_r of A is approximately 0.0365. Thus s was $1.4771 \cdot 10^{-4}$ for $f = 0.1$, 0.0013 for $f = 0.5$ and 0.0120 for $f = 0.9$ in the unweighted case where $D = I$. For the weighted case $\sigma_r \approx 0.0222$ and s is thus $5.4840 \cdot 10^{-5}$, $4.9356 \cdot 10^{-4}$ and 0.0044 for f equal to 0.1, 0.5 and 0.9 respectively. The error graphs in Figure 6.1 plot the Euclidean norm of the difference between the current iterate and the minimum norm solution computed by Matlab. The weighted minimum norm solution can be restated as an unweighted minimum norm solution of a modified system:

$$x_D^* = \underset{A^TAx=A^Tb}{\operatorname{argmin}}\, x^TDx = D^{-\frac{1}{2}} \underset{A^TAD^{-\frac{1}{2}}y=A^Tb}{\operatorname{argmin}}\, y^Ty \tag{85}$$

Hence the weighted minimum norm solution x_D^* can be computed by applying the Moore-Penrose pseudo-inverse of $A^TAD^{-\frac{1}{2}}$ to the right-hand side A^Tb.

$$x_D^* = (A^TAD^{-\frac{1}{2}})^+A^Tb \tag{86}$$

Thus the reference solution was obtained by using the `pinv` command or in the case where $D = I$ by computing $V\Sigma^+U^Tb$ which is evidently the same. This can be verified by inserting the singular value decomposition of A into (86). The SVD was computed by the standard `svd` command in Matlab. The unweighted minimum norm solution obtained had the norm $\sqrt{x^Tx} = 10.9296$ and the weighted minimum norm solution obtained had the norm $\sqrt{x^T D x} = 14.0871$. Though the residual does not make any statements on whether the iterate approaches the minimum norm solution (as opposed to an arbitrary solution) it seems to be a good indicator.

6.2 Test Case: Mobile

Here the system matrix $A \in \mathbb{R}^{1013 \times 570}$ has full column-rank ($\operatorname{rank} A = 570$) and thus $A^T A$ has full rank. The smallest singular value σ_r of A was determined to be approximately 0.1612. Thus the parameter s was chosen to be 0.0029, 0.0260 and 0.2338 for f equal to 0.1, 0.5 and 0.9 respectively. For the weighted case $\sigma_r \approx 0.1010$ of AD^{-1} and s therefore 0.0011, 0.0102 and 0.0917. The (weighted) norms of the solutions were 0.1549 in the unweighted case and 0.1936 in the weighted case. Figure 6.2 shows residual and error graphs for the tests.

6.3 Test Case: Pyramid

In the last test case $A \in \mathbb{R}^{1155 \times 1240}$ has full row-rank ($\operatorname{rank} A = 1115$) and thus $A^T A$ is again rank deficient. The smallest nonzero singular value σ_r of A was determined to be approximately 0.1202. Thus the parameter s was chosen to be 0.0016, 0.0145 and 0.1301 for f equal to 0.1, 0.5 and 0.9 respectively. For the weighted case $\sigma_r \approx 0.0628$ of AD^{-1} and s therefore $4.3839 \cdot 10^{-4}$, 0.0039 and 0.0355. The (weighted) norms of the solutions were 0.6758 in the unweighted case and 0.8973 in the weighted case. Figure 6.3 shows residual and error graphs for the tests.

7 Conclusion

In this report we showed that the iterated regularization scheme due to Riley sometimes also called the iterated Tikhonov regularization can be generalized to damped least squares problems where D is not necessarily the identity but may be SPD. We showed that the iterative scheme approaches the same point as the unique solutions of the regularized problem (45) when $s \to 0$. Furthermore this point can be characterized as the solution of weighted minimum Euclidean norm (57). Finally several numerical experiments were performed in the field of rigid multibody dynamics supporting the theoretical claims. As future work a generalization of the connections to linear complementarity problems is planed. This generalization would be useful for contact problems from rigid multibody dynamics.

References

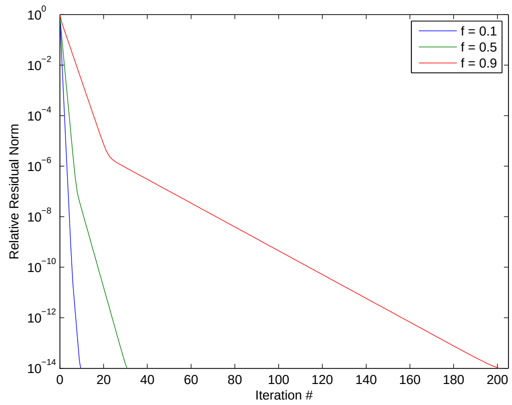
[1] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. Lapack users' guide, 1999.

[2] Ake Björck. *Numerical Methods for Least Squares Problems.* Society for Industrial and Applied Mathematics, 1996.

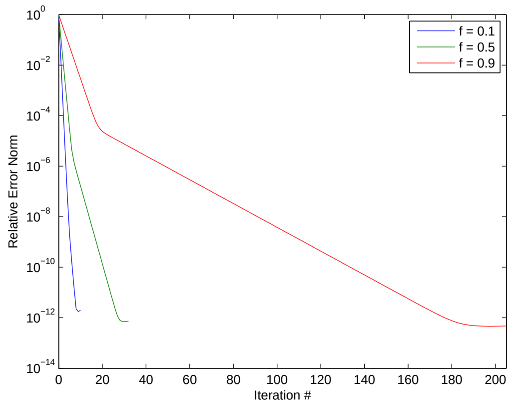
[3] G.H. Golub. Numerical methods for solving linear least squares problems. *Numerische Mathematik*, 7(3):206–216, 1965.

[4] G.H. Golub and C.F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, third edition, 1996.

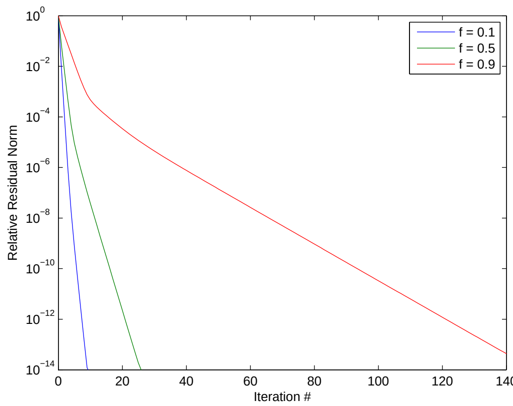
[5] C.L. Lawson and R.J. Hanson. *Solving Least Squares Problems.* Society for Industrial and Applied Mathematics, 1995.



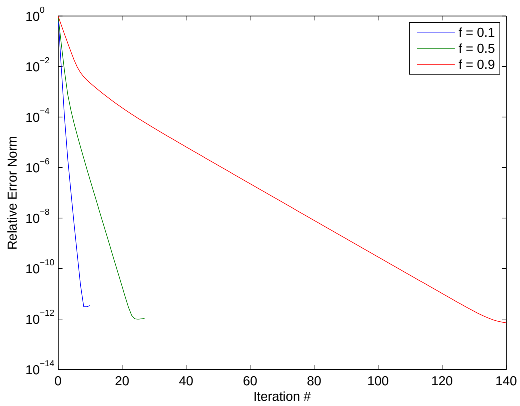
(a) The residual graphs with unweighted regularization.



(b) The error graphs with unweighted regularization.

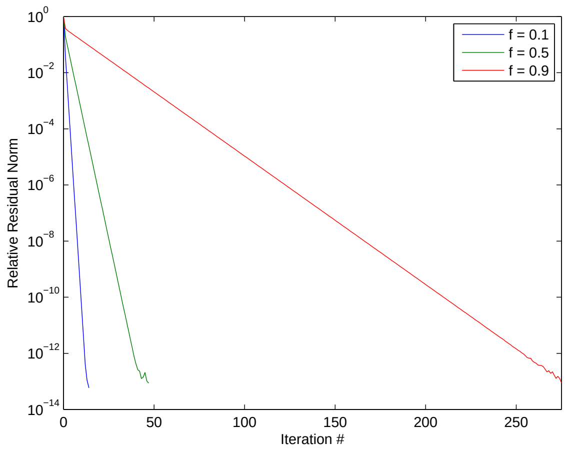


(c) The residual graphs with weighted regularization.

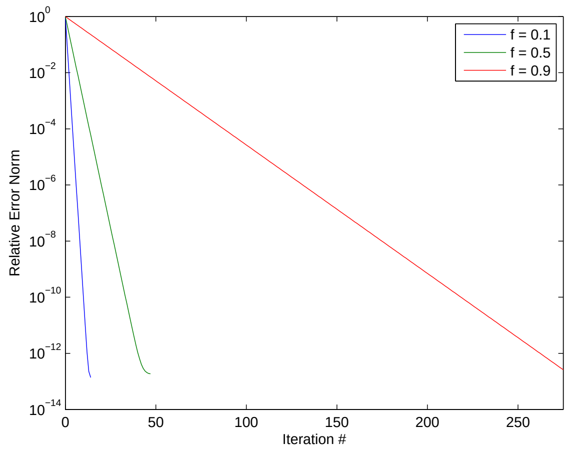


(d) The error graphs with weighted regularization.

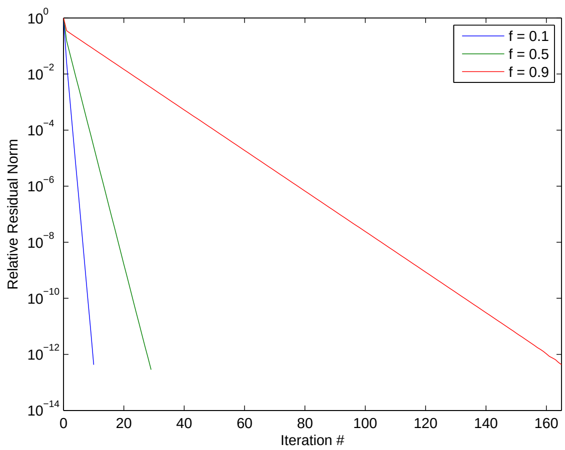
Figure 1: Residual and error graphs for the well test case with different regularizations.



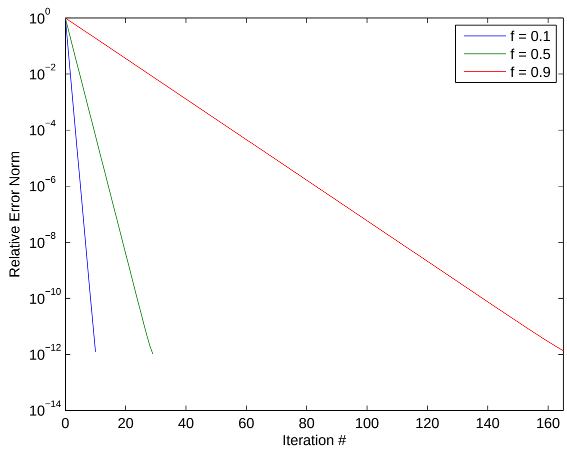
(a) The residual graphs with unweighted regularization.



(b) The error graphs with unweighted regularization.

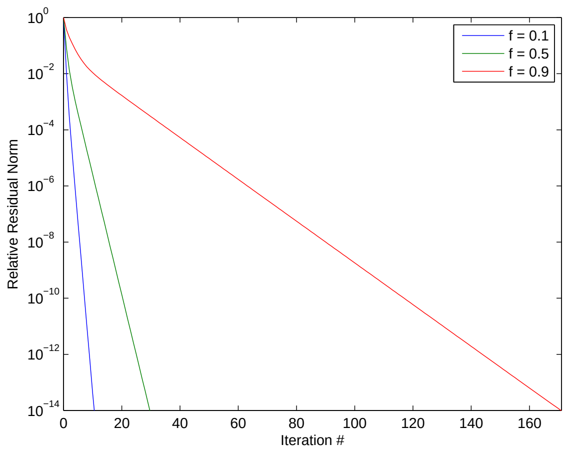


(c) The residual graphs with weighted regularization.

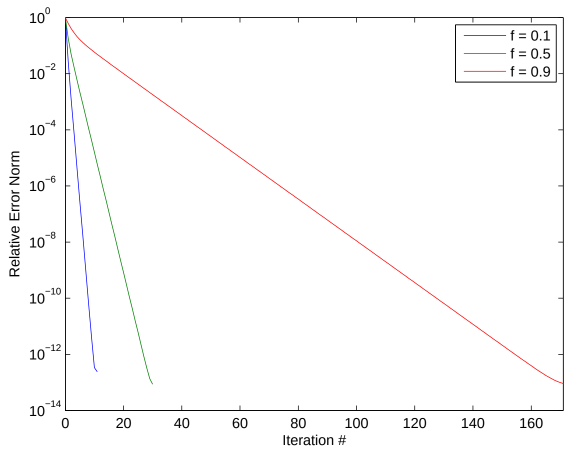


(d) The error graphs with weighted regularization.

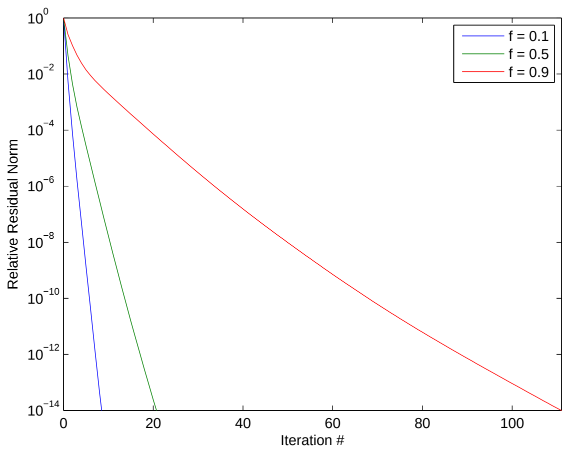
Figure 2: Residual and error graphs for the mobile test case with different regularizations.



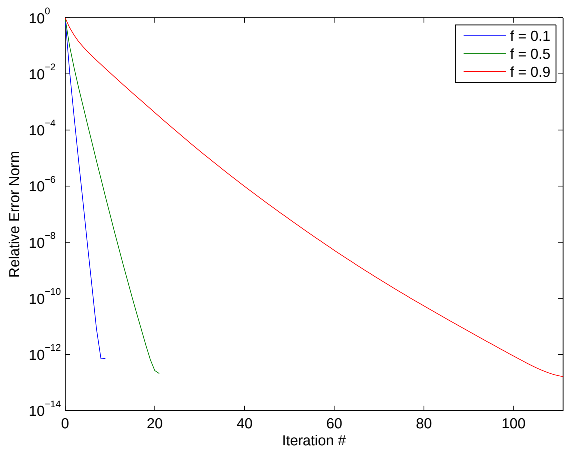
(a) The residual graphs with unweighted regularization.



(b) The error graphs with unweighted regularization.



(c) The residual graphs with weighted regularization.



(d) The error graphs with weighted regularization.

Figure 3: Residual and error graphs for the pyramid test case with different regularizations.

[6] C.F. Van Loan. Generalizing the singular value decomposition. *SIAM Journal on Numerical Analysis*, 13(1):76–83, 1976.

[7] C.C. Paige and M.A. Saunders. Towards a generalized singular value decomposition. *SIAM Journal on Numerical Analysis*, 18(3):398–405, 1981.

[8] C. Popa, T. Preclik, H. Köstler, and U. Rüde. Some projection-based direct solvers for general linear systems of equations. Technical report, Friedrich-Alexander University Erlangen-Nuremberg, 2010.

[9] T. Preclik. Resolving ill-posedness of rigid multibody dynamics. Technical report, Friedrich-Alexander University Erlangen-Nuremberg, 2010. To appear.

[10] J.D. Riley. Solving systems of linear equations with a positive definite, symmetric, but possibly ill-conditioned matrix. *Mathematical Tables and Other Aids to Computation*, 9(51):96–101, 1955.

[11] G.W. Stewart. On the early history of the singular value decomposition. Technical Report 4, 1993.