LBL-37976
UC-405
Preprint

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

Physics Division

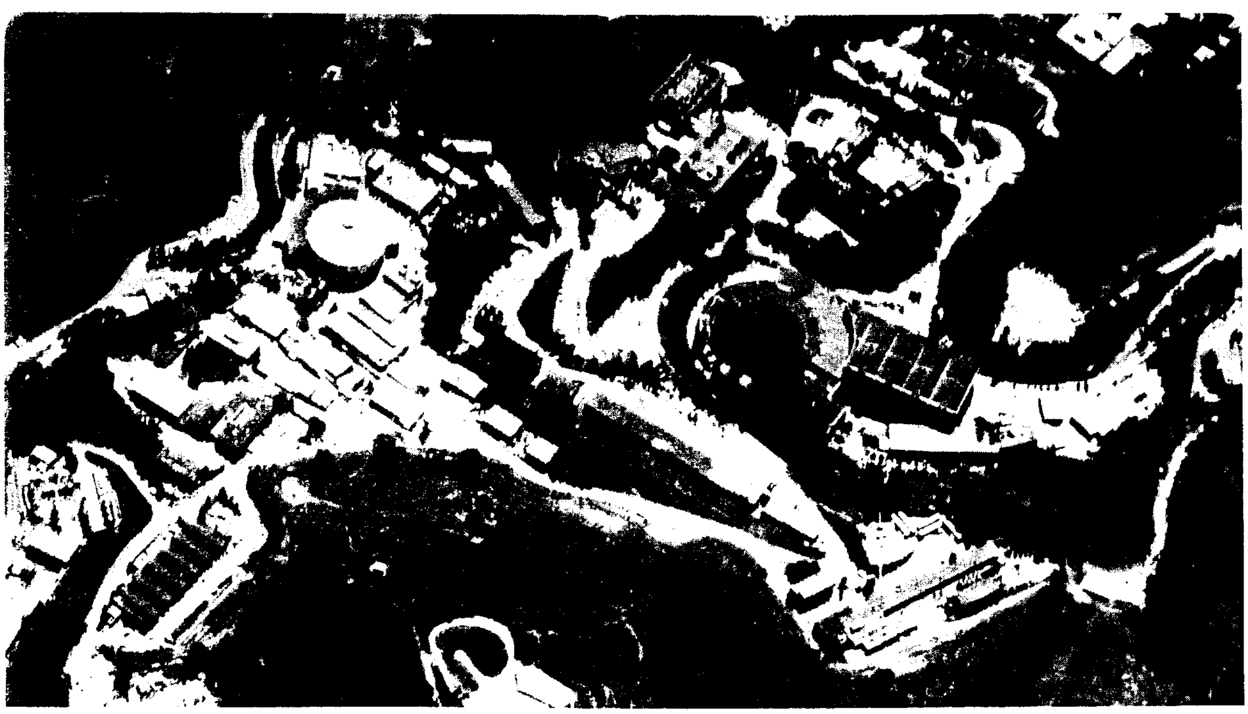
Mathemathics Department

To be submitted for publication

Backward Perturbation Bounds for Linear Least Squares Problems

M. Gu

December 1995



Prepared for the U.S. Department of Energy under Contract Number DE-AC03-76SF00098

LBL-37976

Backward Perturbation Bounds for Linear Least Squares Problems[1]

Ming Gu
Department of Mathematics and Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

December 1995

[1]Supported by the Applied Mathematical Sciences Subprogram of the Office of Energy Research, U.S. Department of Energy under Contract DE-AC03-76SF00098.

Backward Perturbation Bounds for Linear Least Squares Problems

Ming Gu*

December 15, 1995

Abstract

Recently, Higham and Waldén, Karlson, and Sun have provided formulas for computing the smallest backward perturbation bounds for the linear least squares problems. In this paper we provide several backward perturbation bounds that are easier to compute and optimal up to a factor of about 1.6. We also show that any least squares algorithm that is stable in the sense of Stewart is necessarily a backward stable algorithm.

Keywords: Linear least squares problem, perturbation bounds, stability.

*Department of Mathematics and Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720. The author was supported in part by the Applied Mathematical Sciences Subprogram of the Office of Energy Research, U.S. Department of Energy under Contract DE-AC03-76SF00098.

1 Introduction

Given a non-singular matrix $M \in \mathbf{R}^{m\times n}$ with $m \geq n$ and a vector $b \in \mathbf{R}^m$, the linear least squares problem is

$$\min_x \|M \cdot x - b\|_2 , \tag{1.1}$$

which has a unique solution

$$x_M = \left(M^T \cdot M\right)^{-1} \cdot M^T \cdot b .$$

Let $\mathcal{A}$ be an algorithm for solving (1.1) and let $\hat{x}_M \in \mathbf{R}^n$ be the numerical solution computed by $\mathcal{A}$. We say that $\mathcal{A}$ is numerically *stable* if for any such M and b, there exist small perturbation matrices and vectors $\delta M \in \mathbf{R}^{m\times n}$, $\delta b \in \mathbf{R}^m$ and $\delta \hat{x}_M \in \mathbf{R}^n$ such that (see Stewart [6, pages 75-76])

$$\|(M + \delta M) \cdot (\hat{x}_M + \delta \hat{x}_M) - (b + \delta b)\|_2 = \min_x \|(M + \delta M) \cdot x - (b + \delta b)\|_2 . \tag{1.2}$$

It is well-known that if M is ill-conditioned, then $\hat{x}_M$ can be very different from the exact solution x_M (see Higham [4, Chapter 19]). We will call $\hat{x}_M$ a *stable solution* to (1.1) if it satisfies (1.2) for small perturbations δM, δb and $\delta \hat{x}_M$.

The standard method for solving (1.1) via the QR-factorization of M produces a numerical solution $\hat{x}_M$ which satisfies (1.2) with $\delta b = 0$ and $\delta \hat{x}_M = 0$ (see, for example, [4, Chapter 19]). However, if M is a structured matrix such as the Toeplitz matrix, then there are fast algorithms that produce a numerical solution $\hat{x}_M$ which satisfies (1.2) with non-zero δb and $\delta \hat{x}_M$ (see, for example, Gu [2]).

In this paper, we consider the following problem: Given a vector $\hat{x}_M \in \mathbf{R}^n$, find out whether it is a stable solution to (1.1). We will solve this problem by finding out whether there exist small perturbations δM, δb and $\delta \hat{x}_M$, for which $\hat{x}_M + \delta \hat{x}_M$ satisfies (1.2). For simplicity, we assume throughout this paper that $M \neq 0$ and $b \neq 0$.

1.1 Backward Perturbation Bounds

As a special case, we first consider the problem of whether there exists a small perturbation $\widehat{\delta M} \in \mathbf{R}^{m\times n}$ for which

$$\left\|(M + \widehat{\delta M}) \cdot \hat{x}_M - b\right\|_2 = \min_x \|(M + \widehat{\delta M}) \cdot x - b\|_2 . \tag{1.3}$$

We will call $\hat{x}_M$ a *backward stable solution* to (1.1) if it satisfies (1.3) for a small perturbation $\widehat{\delta M}$. A backward stable solution is a stable solution. As mentioned above, the standard method for solving (1.1) via the QR-factorization of M produces a $\hat{x}_M$ that satisfies (1.3).

In general, the matrix $\widehat{\delta M}$ in (1.3) is not uniquely defined. Recently, Waldén, Karlson, and Sun [7] and Higham [4, Chapter 19] [7] have provided a formula for computing the smallest $\|\widehat{\delta M}\|_F$ among all the possible matrices $\widehat{\delta M}$ that satisfy (1.3).

Theorem 1.1 *Let $r = h - M \cdot \hat{x}_M$ and assume that $\hat{x}_M \neq 0$ and $r \neq 0$. Then the optimal norm-wise backward error in F-norm is*

$$\begin{aligned}\mathcal{E}(\hat{x}_M) &:= \min\{\|\widehat{\delta M}\|_F, \text{ where } \widehat{\delta M} \text{ is a solution to (1.3).}\} \\ &= \min\{\eta, \sigma_{\min}([M \quad \eta \cdot C])\},\end{aligned}$$

where

$$\eta = \frac{\|r\|_2}{\|\hat{x}_M\|_2} \quad \text{and} \quad C = I - \frac{r \cdot r^T}{r^T \cdot r},$$

and $\sigma_{\min}([M \quad \eta \cdot C])$ is the smallest singular value of $[M \quad \eta \cdot C]$.

It is obvious that $\mathcal{E}(\hat{x}_M) = 0$ if $r = 0$. Waldén, Karlson, and Sun [7] also show that $\mathcal{E}(0) = \dfrac{\|M^T \cdot h\|_2}{\|h\|_2}$.

According to Theorem 1.1, $\hat{x}_M$ is a backward stable solution (and hence a stable solution) if $\mathcal{E}(\hat{x}_M)$ is small. However, Theorem 1.1 does not say whether $\hat{x}_M$ is a stable solution if $\mathcal{E}(\hat{x}_M)$ is not small. Although Waldén, Karlson, and Sun [7] have also considered perturbations in b, their results do not completely solve the problem of determining whether $\hat{x}_M$ is a stable solution.

Another problem with Theorem 1.1 is that while $\mathcal{E}(\hat{x}_M)$ is optimal, it is not very straightforward to compute for large m. Since η can be very large for $\hat{x}_M \approx 0$, there could be some numerical difficulty in computing $\mathcal{E}(\hat{x}_M)$ accurately as well.

1.2 Main Results

We provide an alternative F-norm bound on $\widehat{\delta M}$ that is easier to compute and that differs from $\mathcal{E}(\hat{x}_M)$ by at most a factor of about 1.6. Using this bound, we further show that a stable solution in the sense of (1.2) is necessarily a backward stable solution in the sense of (1.3). Hence any stable least squares algorithm is necessarily backward stable. And a numerical solution $\hat{x}_M$ is a stable solution in the sense of (1.2) if and only if $\mathcal{E}(\hat{x}_M)$ is small.

In this paper we only discuss real least squares problems. Our results can be easily extended to the complex case.

2 Alternative Backward Perturbation Bounds

In this section, we express our results in terms of the singular value decomposition (SVD) of M. While it is possible to rewrite these results directly in terms of M, the resulting expressions tend to be more complicated.

Let $M = Q \cdot \begin{pmatrix} D \\ 0 \end{pmatrix} \cdot W^T$ be the SVD of M, where $Q \in \mathbf{R}^{m \times m}$ and $W \in \mathbf{R}^{n \times n}$ are orthogonal; and $D \in \mathbf{R}^{n \times n}$ is non-negative diagonal. Rewrite

$$h = Q \cdot \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \quad \text{and} \quad r = h - M \cdot \hat{x}_M = Q \cdot \begin{pmatrix} r_1 \\ r_2 \end{pmatrix},$$

where h_1 and $r_1 = h_1 - D \cdot (W^T \cdot \hat{x}_M) \in \mathbf{R}^n$; and $h_2 = r_2$. It is well-known that $\gamma := \|r_2\|_2 = \|h - M \cdot x_M\|_2$, $h_1 = D \cdot (W^T \cdot x_M)$, and that $r_1 = 0$ if $\hat{x}_M = x_M$.

Theorem 2.1 *Define*

$$\tilde{\sigma} = \sqrt{\frac{r_1^T \cdot D^2 \cdot (D^2 + \eta^2 I)^{-1} \cdot r_1}{\gamma^2/\eta^2 + \eta^2 \cdot r_1^T \cdot (D^2 + \eta^2 I)^{-2} \cdot r_1}}$$

and[1] $\tilde{\mathcal{E}}(\hat{x}_M) = \min(\eta, \tilde{\sigma})$. *Then*

$$1 \leq \frac{\tilde{\mathcal{E}}(\hat{x}_M)}{\mathcal{E}(\hat{x}_M)} \leq \frac{\sqrt{5}+1}{2} .$$

Proof. Theorem 2.1 obviously holds for $\hat{x}_M = 0$. Hence in the following we assume that $\hat{x}_M \neq 0$. By definition, $\sigma_{\min}([M \quad \eta \cdot C])$ is the smallest non-negative σ such that

$$f(\sigma) := \det\Big(([M \quad \eta \cdot C]) \cdot ([M \quad \eta \cdot C])^T - \sigma^2 I\Big) = 0.$$

Replacing M by its singular value decomposition and simplifying,

$$\begin{aligned}
f(\sigma) &= \det\left(M \cdot M^T + \eta \cdot C^2 - \sigma^2 I\right) = \det\left(M \cdot M^T + (\eta^2 - \sigma^2) \cdot I - \eta^2 \cdot \frac{r \cdot r^T}{\|r\|_2^2}\right) \\
&= \left(\eta^2 - \sigma^2\right)^{m-n-1} \cdot \det\left(\begin{pmatrix} D^2 & 0 \\ 0 & 0 \end{pmatrix} + (\eta^2 - \sigma^2) \cdot I - \frac{\eta^2}{\|r\|_2^2} \cdot \begin{pmatrix} r_1 \\ \gamma \end{pmatrix} \cdot \begin{pmatrix} r_1^T & \gamma \end{pmatrix}\right) \\
&= \left(\eta^2 - \sigma^2\right)^{m-n} \cdot \det\left(D^2 + (\eta^2 - \sigma^2) \cdot I\right) \\
&\quad \cdot \left(1 - \frac{\eta^2 \cdot \gamma^2}{\|r\|_2^2 \cdot (\eta^2 - \sigma^2)} - \frac{\eta^2}{\|r\|_2^2} \cdot r_1^T \cdot \left(D^2 + (\eta^2 - \sigma^2) \cdot I\right)^{-1} \cdot r_1\right) .
\end{aligned}$$

Hence $\sigma_{\min}([M \quad \eta \cdot C])$ is the smallest non-negative $\sigma < \eta$ such that

$$1 - \frac{\eta^2 \cdot \gamma^2}{\|r\|_2^2 \cdot (\eta^2 - \sigma^2)} - \frac{\eta^2}{\|r\|_2^2} \cdot r_1^T \cdot \left(D^2 + (\eta^2 - \sigma^2) \cdot I\right)^{-1} \cdot r_1 = 0 . \tag{2.4}$$

This equation can be rewritten as

$$\begin{aligned}
& 1 - \frac{\eta^2 \cdot \gamma^2}{\|r\|_2^2 \cdot \eta^2} - \frac{\eta^2}{\|r\|_2^2} \cdot r_1^T \cdot \left(D^2 + \eta^2 \cdot I\right)^{-1} \cdot r_1 \\
= \; & \frac{\eta^2 \cdot \gamma^2}{\|r\|_2^2 \cdot (\eta^2 - \sigma^2)} - \frac{\eta^2 \cdot \gamma^2}{\|r\|_2^2 \cdot \eta^2} + \frac{\eta^2}{\|r\|_2^2} \cdot r_1^T \cdot \left(D^2 + (\eta^2 - \sigma^2) \cdot I\right)^{-1} \cdot r_1 \\
& - \frac{\eta^2}{\|r\|_2^2} \cdot r_1^T \cdot \left(D^2 + \eta^2 \cdot I\right)^{-1} \cdot r_1 .
\end{aligned}$$

---

[1] We also define

$$\tilde{\mathcal{E}}(0) = \lim_{\hat{x} \to 0} \tilde{\mathcal{E}}(\hat{x}) = \frac{\|D \cdot r_1\|_2}{\|r\|_2} = \frac{\|M^T \cdot h\|_2}{\|h\|_2} .$$

Since $\|r\|_2^2 = \|r_1\|_2^2 + \gamma^2$, the above equation can be simplified, after some algebra, into

$$\sigma^2 = \frac{r_1^T \cdot D^2 \cdot (D^2 + \eta^2 I)^{-1} \cdot r_1}{\gamma^2/(\eta^2 - \sigma^2) + \eta^2 \cdot r_1^T \cdot (D^2 + (\eta^2 - \sigma^2)I)^{-1} \cdot (D^2 + \eta^2 I)^{-1} \cdot r_1} . \tag{2.5}$$

We note that the expression on the right hand side is $\tilde{\sigma}^2$ if $\sigma = 0$. Since $\gamma^2/\eta^2 \le \gamma^2/(\eta^2 - \sigma^2)$ and

$$r_1^T \cdot \left(D^2 + \eta^2 I\right)^{-2} \cdot r_1 \le r_1^T \cdot \left(D^2 + (\eta^2 - \sigma^2)I\right)^{-1} \cdot \left(D^2 + \eta^2 I\right)^{-1} \cdot r_1$$

for $\sigma \le \eta$, equation (2.5) implies that $\sigma_{\min}([M \quad \eta \cdot C]) \le \tilde{\sigma}$. It follows that $\mathcal{E}(\hat{x}_M) \le \tilde{\mathcal{E}}(\hat{x}_M)$.

We now assume that $\tilde{\sigma} > \eta$. In this case we have $\tilde{\mathcal{E}}(\hat{x}_M) = \eta$. We claim that

$$\sigma_{\min}([M \quad \eta \cdot C]) \ge \beta \cdot \eta \quad \text{where} \quad \beta = \frac{\sqrt{5} - 1}{2} . \tag{2.6}$$

We show this by contradiction. Assume that this was false, so that $\sigma_{\min}([M \quad \eta \cdot C]) < \beta \cdot \eta$. We note that $\gamma^2/\eta^2 > (1 - \beta^2) \cdot \gamma^2/(\eta^2 - \sigma^2)$ and that

$$r_1^T \cdot \left(D^2 + \eta^2 I\right)^{-2} \cdot r_1 > (1 - \beta^2) \cdot r_1^T \cdot \left(D^2 + (\eta^2 - \sigma^2)I\right)^{-1} \cdot \left(D^2 + \eta^2 I\right)^{-1} \cdot r_1$$

for $\sigma < \beta \cdot \eta$. Equation (2.5) now implies that

$$\tilde{\sigma} < \frac{\sigma_{\min}([M \quad \eta \cdot C])}{1 - \beta^2} < \frac{\beta \cdot \eta}{1 - \beta^2} = \eta ,$$

which is a contradiction. Hence relation (2.6) is indeed valid and we have

$$\beta \cdot \eta \le \mathcal{E}(\hat{x}_M) \le \eta .$$

So Theorem 2.1 holds in this case.

We further consider the case where $\tilde{\sigma} \le \eta$. In this case we have $\tilde{\mathcal{E}}(\hat{x}_M) = \tilde{\sigma}$ and $\beta := \sigma_{\min}([M \quad \eta \cdot C])/\tilde{\sigma} \le 1$. Similar to above we have

$$\tilde{\sigma} \le \frac{\sigma_{\min}([M \quad \eta \cdot C])}{1 - \beta^2} \le \frac{\tilde{\sigma} \cdot \beta}{1 - \beta^2} ,$$

which simplifies to

$$1 - \beta^2 \le \beta \quad \text{or} \quad \beta \ge \frac{\sqrt{5} - 1}{2} .$$

It follows that

$$\frac{\sqrt{5} - 1}{2} \cdot \tilde{\sigma} \le \mathcal{E}(\hat{x}_M) \le \tilde{\sigma} .$$

So Theorem 2.1 holds in this case as well. ∎

Hence $\tilde{\mathcal{E}}(\hat{x}_M)$ differs from the smallest possible backward perturbation $\mathcal{E}(\hat{x}_M)$ by a factor of at most $\frac{\sqrt{5}+1}{2} \approx 1.6$. To compute $\tilde{\mathcal{E}}(\hat{x}_M)$, we only need to compute D (the singular values of M) and $Q^T \cdot r$; neither Q nor W need be explicitly computed. This computation can be done, for example, by using the subroutines xGESVD in LAPACK [1].

Equation (2.4) provides an efficient way to compute $\sigma_{\min}([M \quad \eta \cdot C])$ (and hence $\mathcal{E}(\hat{x}_M)$) as well. In fact, equation (2.4) is similar to the *secular equations* solved in Gu and Eisenstat [3] and Li [5]; and their methods can be easily modified to compute $\sigma_{\min}([M \quad \eta \cdot C])$.

In the rest of this section we analyze $\mathcal{E}(\hat{x}_M)$ for several special cases.

Corollary 2.1 *Assume that* $\|r_1\|_2 \le \alpha \cdot \gamma$. *Define*

$$\tilde{\sigma}_1 = \frac{\sqrt{r_1^T \cdot D^2 \cdot (D^2 + \eta^2 I)^{-1} \cdot r_1}}{\|\hat{x}_M\|_2} .$$

Then

$$\frac{1}{\sqrt{1+\alpha^2}} \le \frac{\tilde{\sigma}_1}{\mathcal{E}(\hat{x}_M)} \le \frac{\sqrt{5}+1}{2} .$$

Proof. Since $\|r\|_2^2 = \|r_1\|_2^2 + \gamma^2$ and $\eta^2 = \|r\|_2^2/\|x\|_2^2$, the assumption implies that

$$\begin{aligned} \frac{\|\hat{x}_M\|_2^2}{1+\alpha^2} &\le \gamma^2/\eta^2 \le \gamma^2/\eta^2 + \eta^2 \cdot r_1^T \cdot \left(D^2 + \eta^2 I\right)^{-2} \cdot r_1 \\ &\le \gamma^2/\eta^2 + \eta^2 \cdot r_1^T \cdot r_1/\eta^4 = \frac{\|r\|_2^2}{\eta^2} \\ &= \|\hat{x}_M\|_2^2 . \end{aligned}$$

We also have

$$\frac{\tilde{\sigma}_1}{\tilde{\sigma}} = \frac{\sqrt{\gamma^2/\eta^2 + \eta^2 \cdot r_1^T \cdot (D^2 + \eta^2 I)^{-2} \cdot r_1}}{\|\hat{x}_M\|_2} .$$

Consequently,

$$\frac{1}{\sqrt{1+\alpha^2}} \le \frac{\tilde{\sigma}_1}{\tilde{\sigma}} \le 1 .$$

Corollary 2.1 follows by combining the above relations with Theorem 2.1 and the fact that $\tilde{\sigma}_1 \le \eta$. ∎

The least squares problem (1.1) has a small residual if $\gamma = \|h - M \cdot x_M\|_2 \approx 0$ and large residual otherwise; and $\hat{x}_M = x_M$ if and only if $r_1 = 0$. Since a good approximate solution $\hat{x}_M$ always makes r_1 small, Corollary 2.1 implies that for large residual problems $\hat{x}_M$ is a backward stable solution if and only if $\tilde{\sigma}_1$ is small.

Corollary 2.2 below gives a backward perturbation bound for small residual problems.

Corollary 2.2 *Assume that*

$$\|r_1\|_2 \ge \alpha \cdot \gamma \quad \textit{and} \quad \eta \le \sigma_{\min}(M) .$$

Then

$$\frac{\sqrt{5}-1}{2} \cdot \sqrt{\frac{2\alpha^2}{4+\alpha^2}} \cdot \eta \le \mathcal{E}(\hat{x}_M) \le \eta .$$

Proof. Let $\beta = 2\alpha^2/(4+\alpha^2)$ be a scalar. Then

$$\begin{aligned}\tilde{\sigma}^2 - \beta\cdot\eta^2 &= \frac{r_1^T\cdot D^2\cdot(D^2+\eta^2 I)^{-1}\cdot r_1}{\gamma^2/\eta^2+\eta^2\cdot r_1^T\cdot(D^2+\eta^2 I)^{-2}\cdot r_1} - \beta\cdot\eta^2 \\ &= \frac{r_1^T\cdot D^2\cdot(D^2+\eta^2 I)^{-1}\cdot r_1 - \beta\cdot\gamma^2 - \beta\cdot\eta^4\cdot r_1^T\cdot(D^2+\eta^2 I)^{-2}\cdot r_1}{\gamma^2/\eta^2+\eta^2\cdot r_1^T\cdot(D^2+\eta^2 I)^{-2}\cdot r_1}\,.\end{aligned}$$

Since $\eta \le \sigma_{\min}(M) = \sigma_{\min}(D)$, we have

$$r_1^T\cdot D^2\cdot\left(D^2+\eta^2 I\right)^{-1}\cdot r_1 \ge \|r_1\|_2^2/2 \quad\text{and}\quad \eta^4\cdot r_1^T\cdot\left(D^2+\eta^2 I\right)^{-2}\cdot r_1 \le \|r_1\|_2^2/4\,.$$

Combining these relations and simplifying,

$$\tilde{\sigma}^2-\beta\cdot\eta^2 \ge \frac{\|r_1\|_2^2/2-\beta\cdot\gamma^2-\beta\cdot\|r_1\|_2^2/4}{\gamma^2/\eta^2+\eta^2\cdot r_1^T\cdot(D^2+\eta^2 I)^{-2}\cdot r_1} \ge \frac{\alpha^2\cdot\gamma^2/2-\beta\cdot\gamma^2-\beta\cdot\alpha^2\cdot\gamma^2/4}{\gamma^2/\eta^2+\eta^2\cdot r_1^T\cdot(D^2+\eta^2 I)^{-2}\cdot r_1} = 0\,.$$

It follows that $\tilde{\sigma}^2 \ge \beta\cdot\eta^2$ and that

$$\sqrt{\beta}\cdot\eta \le \tilde{\mathcal{E}}\left(\hat{x}_M\right) \le \eta\,.$$

Corollary 2.2 follows by combining this relation with Theorem 2.1. ∎

3 A Stable Solution is a Backward Stable Solution

In this section we show that a stable solution in the sense of (1.2) is a backward stable solution in the sense of (1.3).

Theorem 3.1 *In (1.2) let δM, δb and $\delta\hat{x}_M$ be small perturbations of M, b, and $\hat{x}_M$, respectively. Then there exists a matrix $\widehat{\delta M} \in \mathbf{R}^{m\times n}$ satisfying (1.3) with*[2]

$$\frac{\|\widehat{\delta M}\|_2}{\|M\|_2} \le \frac{\|\delta M\|_2}{\|M\|_2} + 2\left(1+\frac{\|\delta M\|_2}{\|M\|_2}\right)\cdot\left(\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\|\delta h\|_2}{\|h\|_2}\right)\Big/\left(1-2\cdot\frac{\|\delta h\|_2}{\|h\|_2}\right)\,.$$

Proof. We prove this theorem by applying backward perturbation bounds in §2 to $M+\delta M$. Let the SVD of $M+\delta M$ be $\hat{Q}\cdot\begin{pmatrix}\hat{D}\\0\end{pmatrix}\cdot\hat{W}^T$. Define

$$\hat{Q}^T\cdot h = \begin{pmatrix}\hat{h}_1\\ \hat{h}_2\end{pmatrix}\,,\quad \delta\hat{h} = \hat{Q}^T\cdot\delta h = \begin{pmatrix}\delta\hat{h}_1\\ \delta\hat{h}_2\end{pmatrix} \quad\text{and}\quad \widehat{\delta x}_M = \hat{W}^T\cdot\delta\hat{x}_M\,,$$

where $\hat{h}_1$ and $\delta\hat{h}_1 \in \mathbf{R}^n$; and $\hat{h}_2$ and $\delta\hat{h}_2 \in \mathbf{R}^{(m-n)}$. It follows that $\|\delta\hat{h}\|_2 = \|\delta h\|_2$, $\|\widehat{\delta x}_M\|_2 = \|\delta\hat{x}_M\|_2$. Write

$$\begin{aligned}\hat{r} &:= h-(M+\delta M)\cdot\hat{x}_M \\ &= ((h+\delta h)-(M+\delta M)\cdot(\hat{x}_M+\delta\hat{x}_M))+((M+\delta M)\cdot\delta\hat{x}_M-\delta h)\,.\end{aligned} \tag{3.7}$$

[2]In the event $\hat{x}_M = 0$, we adopt the convention that $\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2} = 0$ if $\delta\hat{x}_M = 0$.

Since $\hat{x}_M + \delta\hat{x}_M$ is the exact solution to the perturbed least squares problem (1.2), we have

$$(h + \delta h) - (M + \delta M) \cdot (\hat{x}_M + \delta\hat{x}_M) = \hat{Q} \cdot \begin{pmatrix} 0 \\ \hat{h}_2 + \delta\hat{h}_2 \end{pmatrix} ,$$

We also have

$$(M + \delta M) \cdot \delta\hat{x}_M - \delta h = \hat{Q} \cdot \begin{pmatrix} \hat{D} \cdot \widehat{\delta x}_M - \delta\hat{h}_1 \\ -\delta\hat{h}_2 \end{pmatrix} .$$

Plugging these relations into (3.7) we have

$$\hat{r} = \hat{Q} \cdot \begin{pmatrix} \hat{D} \cdot \widehat{\delta x}_M - \delta\hat{h}_1 \\ \hat{h}_2 \end{pmatrix} .$$

In the following we derive an upper bound on $\mathcal{E}(\hat{x}_M)$ with $M + \delta M$ as the coefficient matrix in the least squares problem. Define

$$\hat{\eta} = \frac{\|\hat{r}\|_2}{\|\hat{x}_M\|_2} , \quad \hat{\gamma} = \|\hat{h}_2\|_2 \quad \text{and} \quad \hat{r}_1 = \hat{D} \cdot \widehat{\delta x}_M - \delta\hat{h}_1 .$$

We first assume that $\hat{\gamma} \leq \|\hat{r}_1\|_2$. By Theorem 2.1,

$$\begin{aligned} \mathcal{E}(\hat{x}_M) \quad &\leq \quad \hat{\eta} = \frac{\|\hat{r}\|_2}{\|\hat{x}_M\|_2} \leq \sqrt{2} \cdot \frac{\|\hat{r}_1\|_2}{\|\hat{x}_M\|_2} \leq \sqrt{2} \cdot \frac{\|\hat{D} \cdot \widehat{\delta x}_M\|_2 + \|\delta\hat{h}_1\|_2}{\|\hat{x}_M\|_2} \\ &\leq \quad \sqrt{2} \cdot \left(\|\hat{D}\|_2 \cdot \frac{\|\widehat{\delta x}_M\|_2}{\|\hat{x}_M\|_2} + \frac{\|\delta\hat{h}_1\|_2}{\|\hat{x}_M\|_2} \right) . \end{aligned} \tag{3.8}$$

Since $\hat{x}_M + \delta\hat{x}_M$ is the exact solution to (1.2), it follows that

$$\hat{D} \cdot W^T \cdot (\hat{x}_M + \delta\hat{x}_M) = \hat{h}_1 + \delta\hat{h}_1 ,$$

and hence

$$\|\hat{h}_1 + \delta\hat{h}_1\|_2 \leq \|\hat{D}\|_2 \cdot \|\hat{x}_M + \delta\hat{x}_M\|_2 \leq \|\hat{D}\|_2 \cdot \|\hat{x}_M\|_2 \cdot \left(1 + \frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}\right) . \tag{3.9}$$

On the other hand,

$$\hat{Q}^T \cdot h = \begin{pmatrix} \hat{h}_1 \\ \hat{h}_2 \end{pmatrix} = \begin{pmatrix} \hat{h}_1 + \delta\hat{h}_1 \\ 0 \end{pmatrix} + \begin{pmatrix} -\delta\hat{h}_1 \\ \hat{h}_2 \end{pmatrix} .$$

Taking 2-norms on both sides,

$$\begin{aligned} \|h\|_2 \quad &\leq \quad \left\|\hat{h}_1 + \delta\hat{h}_1\right\|_2 + \left\|\delta\hat{h}_1\right\|_2 + \left\|\hat{h}_2\right\|_2 \leq \left\|\hat{h}_1 + \delta\hat{h}_1\right\|_2 + \|\delta h\|_2 + \|\hat{r}_1\|_2 \\ &= \quad \left\|\hat{h}_1 + \delta\hat{h}_1\right\|_2 + \|\delta h\|_2 + \left\|\hat{D} \cdot \widehat{\delta x}_M - \delta\hat{h}_1\right\|_2 . \end{aligned}$$

Plugging in the 2-norm upper bound on $\|\hat{h}_1 + \delta\hat{h}_1\|_2$ and simplifying, we get

$$\frac{\|h\|_2}{\|\hat{x}_M\|_2} \leq \|\hat{D}\|_2 \cdot \left(1 + 2 \cdot \frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}\right) \Big/ \left(1 - 2 \cdot \frac{\|\delta h\|_2}{\|h\|_2}\right) .$$

In (3.8) we have

$$\begin{aligned}\mathcal{E}(\hat{x}_M) &\le \sqrt{2}\cdot\left(\|\hat{D}\|_2\cdot\frac{\|\widehat{\delta x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\|\delta\hat{h}_1\|_2}{\|h\|_2}\cdot\frac{\|h\|_2}{\|\hat{x}_M\|_2}\right)\\ &\le \sqrt{2}\cdot\left(\|\hat{D}\|_2\cdot\frac{\|\widehat{\delta x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\|\delta h\|_2}{\|h\|_2}\cdot\frac{\|h\|_2}{\|\hat{x}_M\|_2}\right).\end{aligned}$$

Plugging in the upper bound on $\frac{\|h\|_2}{\|\hat{x}_M\|_2}$ and simplifying, we obtain

$$\begin{aligned}\mathcal{E}(\hat{x}_M) &\le \sqrt{2}\cdot(\|M\|_2+\|\delta M\|_2)\cdot\left(\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\|\delta h\|_2}{\|h\|_2}\right)\Big/\left(1-2\cdot\frac{\|\delta h\|_2}{\|h\|_2}\right)\\ &\le 2\cdot(\|M\|_2+\|\delta M\|_2)\cdot\left(\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\|\delta h\|_2}{\|h\|_2}\right)\Big/\left(1-2\cdot\frac{\|\delta h\|_2}{\|h\|_2}\right), \qquad (3.10)\end{aligned}$$

where we have used the fact that

$$\|\hat{D}\|_2=\|M+\delta M\|_2\le\|M\|_2+\|\delta M\|_2 .$$

Now we assume that $\hat{\gamma}\ge\|\hat{r}_1\|_2$. By Corollary 2.1 we have $\mathcal{E}(\hat{x}_M)\le\sqrt{2}\cdot\tilde{\sigma}_1$, where

$$\tilde{\sigma}_1=\frac{\sqrt{\hat{r}_1^T\cdot\hat{D}^2\cdot\left(\hat{D}^2+\hat{\eta}^2 I\right)^{-1}\cdot\hat{r}_1}}{\|\hat{x}_M\|_2}.$$

Since $\hat{r}_1=\hat{D}\cdot\widehat{\delta x}_M-\delta\hat{h}_1$, it follows that

$$\begin{aligned}\tilde{\sigma}_1 &\le \frac{\sqrt{\left(\hat{D}\cdot\widehat{\delta x}_M\right)^T\cdot\hat{D}^2\cdot\left(\hat{D}^2+\hat{\eta}^2 I\right)^{-1}\cdot\left(\hat{D}\cdot\widehat{\delta x}_M\right)}}{\|\hat{x}_M\|_2}\\ &\quad+\frac{\sqrt{\delta\hat{h}_1^T\cdot\hat{D}^2\cdot\left(\hat{D}^2+\hat{\eta}^2 I\right)^{-1}\cdot\delta\hat{h}_1}}{\|\hat{x}_M\|_2}\\ &\le \|\hat{D}\|_2\cdot\frac{\|\widehat{\delta x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\sqrt{\delta\hat{h}_1^T\cdot\hat{D}^2\cdot\left(\hat{D}^2+\hat{\eta}^2 I\right)^{-1}\cdot\delta\hat{h}_1}}{\|\hat{x}_M\|_2}.\end{aligned}$$

Since $\|\hat{r}\|_2=\hat{\eta}\cdot\|\hat{x}_M\|_2$, it follows from the above relation that

$$\tilde{\sigma}_1\le\|\hat{D}\|_2\cdot\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\|\delta\hat{h}_1\|_2}{\|\hat{x}_M\|_2}\quad\text{and}\quad\tilde{\sigma}_1\le\|\hat{D}\|_2\cdot\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\|D\cdot\delta\hat{h}_1\|_2}{\|\hat{r}\|_2}.$$

Combining these with relation (3.9) we obtain

$$\begin{aligned}\tilde{\sigma}_1 &\le \|\hat{D}\|_2\cdot\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}+\min\left(\frac{\|\delta h\|_2}{\|\hat{h}_1+\delta\hat{h}_1\|_2},\frac{\|\delta h\|_2}{\|\hat{r}\|_2}\right)\cdot\|\hat{D}\|_2\cdot\left(1+\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}\right)\\ &= \|\hat{D}\|_2\cdot\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}+\frac{\|\delta h\|_2}{\max\left(\|\hat{h}_1+\delta\hat{h}_1\|_2,\|\hat{r}\|_2\right)}\cdot\|\hat{D}\|_2\cdot\left(1+\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}\right).\end{aligned}$$

Since $\|\hat{r}\|_2 \geq \|\hat{h}_2\|_2$, it follows that

$$\begin{aligned}\max\left(\|\hat{h}_1 + \delta\hat{h}_1\|_2, \|\hat{r}\|_2\right) &\geq \max\left(\|\hat{h}_1 + \delta\hat{h}_1\|_2, \|\hat{h}_2\|_2\right) \geq \frac{1}{\sqrt{2}} \cdot \sqrt{\|\hat{h}_1 + \delta\hat{h}_1\|_2^2 + \|\hat{h}_2\|_2^2} \\ &= \frac{1}{\sqrt{2}} \cdot \left\| \begin{pmatrix} \hat{h}_1 + \delta\hat{h}_1 \\ \hat{h}_2 \end{pmatrix} \right\|_2 \geq \frac{1}{\sqrt{2}} \cdot \left(\|h\|_2 - \|\delta h\|_2\right) .\end{aligned}$$

Consequently,

$$\tilde{\sigma}_1 \leq \|\hat{D}\|_2 \cdot \frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2} + \frac{\sqrt{2} \cdot \|\delta h\|_2}{\|h\|_2 - \|\delta h\|_2} \cdot \|\hat{D}\|_2 \cdot \left(1 + \frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2}\right) .$$

From this relation we get

$$\begin{aligned}\mathcal{E}\left(\hat{x}_M\right) &\leq 2\left(\|M\|_2 + \|\delta M\|_2\right) \cdot \left(\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2} + \frac{\|\delta h\|_2}{\|h\|_2}\right) \Big/ \left(1 - \frac{\|\delta h\|_2}{\|h\|_2}\right) \\ &\leq 2\left(\|M\|_2 + \|\delta M\|_2\right) \cdot \left(\frac{\|\delta\hat{x}_M\|_2}{\|\hat{x}_M\|_2} + \frac{\|\delta h\|_2}{\|h\|_2}\right) \Big/ \left(1 - 2 \cdot \frac{\|\delta h\|_2}{\|h\|_2}\right) ,\end{aligned}$$

which is identical to (3.10).

In both cases, there exists a matrix $\widehat{\delta M}_1 \in \mathbf{R}^{m \times n}$ with $\|\widehat{\delta M}_1\|_F = \mathcal{E}\left(\hat{x}_M\right)$ such that

$$\|(M + \delta M + \widehat{\delta M}_1) \cdot \hat{x}_M - h\|_2 = \min_x \|(M + \delta M + \widehat{\delta M}_1) \cdot x - h\|_2 .$$

Now we define $\widehat{\delta M} = \delta M + \widehat{\delta M}_1$. It follows that

$$\|(M + \widehat{\delta M}) \cdot \hat{x}_M - h\|_2 = \min_x \|(M + \widehat{\delta M}) \cdot x - h\|_2 ,$$

and that

$$\|\widehat{\delta M}\|_2 \leq \|\delta M\|_2 + \mathcal{E}\left(\hat{x}_M\right) .$$

The theorem follows immediately by plugging the upper bound (3.10) into this relation. ∎

References

[1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen. *LAPACK Users' Guide.* SIAM, Philadelphia, PA, second edition, 1994.

[2] M. Gu. New fast algorithms for structured linear least squares problems. LBL Report LBL-37878, Lawrence Berkeley National Laboratory, November 1995.

[3] M. Gu and S. C. Eisenstat. A divide-and-conquer algorithm for the bidiagonal SVD. *SIAM J. Matrix Anal. Appl.*, 16:79–92, 1995.

[4] N. J. Higham. *Accuracy and Stability of Numerical Algorithms.* SIAM, 1996. To appear.

[5] R.-C. Li. Solving secular equations stably and efficiently. Unpublished manuscript, October 1992.

[6] G. W. Stewart. *Introduction to Matrix Computations*. Academic Press, New York, 1973.

[7] B. Waldén, R. Karlson, and J.-G. Sun. Optimal backward perturbation bounds for the linear least square problem. *Numer. Lin. Alg. Appl.*, 2:271–286, 1995.