BIT Numerical Mathematics (2006)
DOI: 10.1007/s10543-006-0049-0
Published online: 6 March 2006 –

CONDITION NUMBERS FOR STRUCTURED LEAST SQUARES PROBLEMS*

W. XU[1], Y. WEI[2] and S. QIAO[1,**]

[1] *Department of Computing and Software, McMaster University Hamilton, Ontario, L8S 4L7, Canada.*
emails: {xuw5,qiao}@mcmaster.ca

[2] *School of Mathematical Sciences, Fudan University Shanghai, 200433, P.R. China.*
Key Laboratory of Mathematics for Nonlinear Sciences (Fudan University), Ministry of Education, Shanghai, 200433, P.R. China.
email: ymwei@fudan.edu.cn

Abstract.

This paper studies the normwise perturbation theory for structured least squares problems. The structures under investigation are symmetric, persymmetric, skewsymmetric, Toeplitz and Hankel. We present the condition numbers for structured least squares.

AMS subject classification (2000): 15A18, 65F20, 65F25, 65F50.

Key words: condition numbers, symmetric matrix, skewsymmetric matrix, Toeplitz matrix, Hankel matrix, circulant matrix and structured matrix.

1 Introduction.

A condition number is a measurement of the sensitivity of a problem to the perturbations in the data. In the problem of solving the linear system

$$A\mathbf{x} = \mathbf{b},$$

where $A \in R^{n\times n}$ and is nonsingular and $\mathbf{b} \in R^n$, the condition number measures the sensitivity of the solution $\mathbf{x}$ to the perturbations in A and $\mathbf{b}$. Suppose that the perturbations in A and $\mathbf{b}$ are ΔA, $\|\Delta A\| \le \epsilon\|A\|$, and $\Delta\mathbf{b}$, $\|\Delta\mathbf{b}\| \le \epsilon\|\mathbf{b}\|$, respectively and $A + \Delta A$ is nonsingular. Note that, in this paper, $\|\cdot\|$ denotes the 2-norm of a vector or the spectral norm of a matrix. If $\mathbf{y}$ is the solution of

* Received March 2003. Accepted in revised form January 2006. Communicated by Per Christian Hansen.

** This work is partially supported by the Key Laboratory of Mathematics for Nonlinear Sciences at Fudan University during Sanzheng Qiao's visit. The second author is supported by the National Natural Science Foundation of China under grant 10471027 and Shanghai Education Committee. The third author is partially supported by Natural Science and Engineering Research Council of Canada under grant OGP0046301.

the perturbed system $(A+\Delta A)\mathbf{y} = \mathbf{b}+\Delta\mathbf{b}$, then

$$\frac{\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}\|} \le \kappa_1 \frac{\epsilon}{1-\epsilon\|A^{-1}\|\,\|A\|}$$

if $\epsilon\|A^{-1}\|\,\|A\| < 1$ [6, p. 120], where the condition number

$$\kappa_1 := \|A^{-1}\|\,\|A\| + \frac{\|A^{-1}\|\,\|\mathbf{b}\|}{\|\mathbf{x}\|} \le 2\|A^{-1}\|\,\|A\|.$$

As shown above, the condition number κ_1 is the linear term in the relative error in the solution $\mathbf{x}$ for small ϵ. Thus we have the following definition of the condition number for a nonsingular linear system [6, p. 121]:

$$\kappa_1 := \lim_{\epsilon\to 0}\sup\left\{\frac{\|\Delta\mathbf{x}\|}{\epsilon\|\mathbf{x}\|} : \quad (A+\Delta A)(\mathbf{x}+\Delta\mathbf{x}) = \mathbf{b}+\Delta\mathbf{b}, \right.$$
$$\left. \|\Delta A\| \le \epsilon\|A\|,\ \|\Delta\mathbf{b}\| \le \epsilon\|\mathbf{b}\| \right\}.$$

When A is structured, for example, symmetric or Toeplitz, it is reasonable to require that the perturbation ΔA have the same structure. Denoting the structure as $\mathcal{S}$, we have the following definition of the condition number for a structured linear system:

$$\kappa_1^{\mathcal{S}} := \lim_{\epsilon\to 0}\sup\left\{\frac{\|\Delta\mathbf{x}\|}{\epsilon\|\mathbf{x}\|} : \quad (A+\Delta A)(\mathbf{x}+\Delta\mathbf{x}) = \mathbf{b}+\Delta\mathbf{b}, \right.$$
$$\left. A, \Delta A \in \mathcal{S},\ \|\Delta A\| \le \epsilon\|A\|,\ \|\Delta\mathbf{b}\| \le \epsilon\|\mathbf{b}\| \right\}. \tag{1.1}$$

Rump [8] presents a perturbation theory for structured linear systems and gives tight upper bounds for the condition numbers for structured matrices such as symmetric, persymmetric, skewsymmetric, circulant, Toeplitz, and Hankel.

In this paper, we consider a general case, the least squares problem:

$$\|A\mathbf{x}-\mathbf{b}\| = \min. \tag{1.2}$$

The least norm least squares solution is $\mathbf{x} = A^+\mathbf{b}$, where A^+ is the Moore–Penrose inverse of A [4, 10], the unique matrix satisfying

$$AA^+A = A, \quad A^+AA^+ = A^+, \quad (AA^+)^{\mathrm{H}} = AA^+, \quad (A^+A)^{\mathrm{H}} = A^+A.$$

Let $\mathbf{y}$ be the least norm solution of the perturbed least squares problem

$$\|(A+\Delta A)\mathbf{y} - (\mathbf{b}+\Delta\mathbf{b})\| = \min,$$

then [6, p. 382]

$$\frac{\|\mathbf{x}-\mathbf{y}\|}{\|\mathbf{x}\|} \le \frac{\epsilon\|A\|\,\|A^+\|}{1-\epsilon\|A\|\,\|A^+\|}\left(2 + (\|A\|\,\|A^+\|+1)\frac{\|A\mathbf{x}-\mathbf{b}\|}{\|A\|\,\|\mathbf{x}\|}\right).$$

To simplify the discussion, we assume the residual $A\mathbf{x} - \mathbf{b}$ is small. Thus the condition number is approximately $2\,\|A\|\,\|A^+\|$. Furthermore, to make the condition number for the least squares analogous to the condition number κ_1 for linear systems, we consider a lower bound

$$\kappa := \|A^+\|\,\|A\| + \|A^+\|\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leq 2\,\|A\|\,\|A^+\|. \tag{1.3}$$

For no perturbation on the right hand side, we define

$$\kappa_0 := \|A^+\|\,\|A\|. \tag{1.4}$$

This paper studies the condition numbers for structured least squares problems. In Section 2, we define the condition number for structured least squares problems, derive bounds for the condition number, and give a general expression of the condition number. We then discuss special structures. The condition number for symmetric, persymmetric, and skewsymmetric least squares problems is discussed in Section 3. Matrices like circulant or Toeplitz have fewer number of independent entries than symmetric matrices. In Section 4, by exploiting the structures, we present a general expression of the condition number for those more structured least squares problems. The specific cases of circulant, symmetric Toeplitz, and Hankel are shown in Sections 5, 6, and 7 respectively. Finally, in Section 8, we present the condition number for the generalized inversion of matrices with structures: symmetric, skewsymmetric, and Hankel.

2 General case.

When the matrix A in the least squares problem (1.2) is structured, it is reasonable to assume that the perturbation ΔA have the same structure. Following the definition (1.1) of the condition number $\kappa_1^{\mathcal{S}}$ for structured linear systems, we define the condition number

$$\begin{aligned}\kappa^{\mathcal{S}} := \lim_{\epsilon\to 0}\sup\Big\{ \frac{\|\Delta\mathbf{x}\|}{\epsilon\|\mathbf{x}\|} :\ & \|(A+\Delta A)(\mathbf{x}+\Delta\mathbf{x}) - (\mathbf{b}+\Delta\mathbf{b})\| = \min,\\ & A, \Delta A \in \mathcal{S},\ \|\Delta A\| \leq \epsilon\|A\|,\ \|\Delta\mathbf{b}\| \leq \epsilon\|\mathbf{b}\| \Big\}\end{aligned} \tag{2.1}$$

for structured least squares problems.

In our analysis, in addition to the structure requirement, we assume that the perturbation ΔA satisfies the conditions:

$$\mathrm{Range}(\Delta A) \subseteq \mathrm{Range}(A) \quad \text{and} \quad \mathrm{Range}((\Delta A)^{\mathrm{T}}) \subseteq \mathrm{Range}(A^{\mathrm{T}}). \tag{2.2}$$

In that sense, our analysis, different from the linear system case where A is assumed nonsingular, is "non-standard". Equivalent to (2.2), $AA^+\Delta A = \Delta A$ and $A^+A(\Delta A)^{\mathrm{T}} = (\Delta A)^{\mathrm{T}}$, then we have [1]

$$(A+\Delta A)^+ = A^+ - A^+(\Delta A)A^+ + O(\epsilon^2). \tag{2.3}$$

Here is an example of two singular matrices A and ΔA satisfying the conditions (2.2) and having the same Toeplitz structure:

$$A = \begin{bmatrix} 1 & 5 & 4 & 3 & 2 & 1 \\ 2 & 1 & 5 & 4 & 3 & 2 \\ 3 & 2 & 1 & 5 & 4 & 3 \\ 4 & 3 & 2 & 1 & 5 & 4 \\ 5 & 4 & 3 & 2 & 1 & 5 \\ 1 & 5 & 4 & 3 & 2 & 1 \end{bmatrix}, \quad \Delta A = \begin{bmatrix} \epsilon_1 & \epsilon_5 & \epsilon_4 & \epsilon_3 & \epsilon_2 & \epsilon_1 \\ \epsilon_2 & \epsilon_1 & \epsilon_5 & \epsilon_4 & \epsilon_3 & \epsilon_2 \\ \epsilon_3 & \epsilon_2 & \epsilon_1 & \epsilon_5 & \epsilon_4 & \epsilon_3 \\ \epsilon_4 & \epsilon_3 & \epsilon_2 & \epsilon_1 & \epsilon_5 & \epsilon_4 \\ \epsilon_5 & \epsilon_4 & \epsilon_3 & \epsilon_2 & \epsilon_1 & \epsilon_5 \\ \epsilon_1 & \epsilon_5 & \epsilon_4 & \epsilon_3 & \epsilon_2 & \epsilon_1 \end{bmatrix},$$

where $\epsilon_1 = 0.153 \times 10^{-4}$, $\epsilon_2 = 7.468 \times 10^{-4}$, $\epsilon_3 = 4.451 \times 10^{-4}$, $\epsilon_4 = 9.318 \times 10^{-4}$, and $\epsilon_5 = 4.660 \times 10^{-4}$.

If ΔA satisfies the conditions (2.2), from (2.3), we get

$$\begin{aligned} \mathbf{x} + \Delta\mathbf{x} &= (A + \Delta A)^+(\mathbf{b} + \Delta\mathbf{b}) \\ &= \mathbf{x} + A^+\Delta\mathbf{b} - A^+(\Delta A)A^+\mathbf{b} + O(\epsilon^2). \end{aligned}$$

It then follows that

$$\Delta\mathbf{x} \approx -A^+(\Delta A\mathbf{x} - \Delta\mathbf{b}). \tag{2.4}$$

In particular, when

$$\Delta\mathbf{b} = -\frac{\|\mathbf{b}\|}{\|A\|\,\|\mathbf{x}\|}\Delta A\,\mathbf{x}, \tag{2.5}$$

then (2.4) turns into

$$\Delta\mathbf{x} \approx -A^+\Delta A\,\mathbf{x}\left(1 + \frac{\|\mathbf{b}\|}{\|A\|\,\|\mathbf{x}\|}\right). \tag{2.6}$$

Substituting $\Delta\mathbf{x}$ in (2.1) with (2.6) for the particular $\Delta\mathbf{b}$ in (2.5), we get

$$\kappa^{\mathcal{S}} \geq \frac{\|A^+\Delta A\,\mathbf{x}\|}{\epsilon\|A\|\,\|\mathbf{x}\|}\left(\|A\| + \frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}\right). \tag{2.7}$$

This motivates us to define a matrix E in $\Delta A = \epsilon\|A\|\,E$, i.e., the linear term in the relative error in A, and a quantity

$$\varphi := \sup\{\|A^+E\mathbf{x}\| :\ \|A\mathbf{x} - \mathbf{b}\| = \min,\ A, E \in \mathcal{S},\ \|E\| \leq 1\}, \tag{2.8}$$

provided that E satisfies the conditions

$$\mathrm{Range}(E) \subseteq \mathrm{Range}(A) \quad \text{and} \quad \mathrm{Range}(E^{\mathrm{T}}) \subseteq \mathrm{Range}(A^{\mathrm{T}}), \tag{2.9}$$

which implies that ΔA satisfies the conditions (2.2). Thus, combining (2.7) and (2.8), we have a lower bound

$$\kappa^{\mathcal{S}} \geq \frac{\varphi}{\|\mathbf{x}\|}\left(\|A\| + \frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}\right).$$

On the other hand, it follows from (2.1) and (2.4) that

$$\kappa^{\mathcal{S}} \leq \|A^{+}\|\,\|A\| + \|A^{+}\|\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}.$$

Putting two bounds together, we have

$$\frac{\varphi}{\|\mathbf{x}\|}\left(\|A\| + \frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}\right) \leq \kappa^{\mathcal{S}} \leq \|A^{+}\|\,\|A\| + \|A^{+}\|\,\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}. \tag{2.10}$$

Especially, when $\varphi = \|A^{+}\|\,\|\mathbf{x}\|$, the equalities in (2.10) can be attained, that is

$$\kappa^{\mathcal{S}} = \|A^{+}\|\,\|A\| + \|A^{+}\|\,\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|} \approx \|A^{+}\|\,\|A\| + \|A^{+}\|\,\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}.$$

Since (1.3) is a lower bound for κ, the ratio $\kappa^{\mathcal{S}}/\kappa$ is always approximately less than or equal to 1. In the following, we construct a matrix E in (2.8) so that $\varphi = \|A^{+}\|\,\|\mathbf{x}\|$.

Suppose that

$$A = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^{\mathrm{T}}$$

is the singular value decomposition (SVD) of A and $\mathrm{rank}(A) = r$. We construct $E = \mathbf{y}\mathbf{x}^{\mathrm{T}}/\|\mathbf{x}\|$, where $\mathbf{x} = A^{+}\mathbf{b}$ and $\mathbf{y} = \mathbf{u}_r$, the rth column of U. Obviously, $\|E\| \leq 1$. Then

$$\|A^{+}\mathbf{y}\| = \left\| V \begin{bmatrix} D^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^{\mathrm{T}}\mathbf{u}_r \right\| = \|A^{+}\|.$$

Consequently,

$$\varphi = \sup \|A^{+}E\mathbf{x}\| = \sup \|A^{+}\mathbf{y}\|\,\|\mathbf{x}\| = \|A^{+}\|\,\|\mathbf{x}\|.$$

The two conditions (2.9) can be verified by

$$AA^{+}E = U \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} U^{\mathrm{T}}\mathbf{u}_r\mathbf{x}^{\mathrm{T}}/\|\mathbf{x}\| = \mathbf{u_r}\mathbf{x}^{\mathrm{T}}/\|\mathbf{x}\| = E$$

and

$$A^{+}A\ E^{\mathrm{T}} = A^{+}A\mathbf{x}\mathbf{u}_r^{\mathrm{T}}/\|\mathbf{x}\| = A^{+}AA^{+}\mathbf{b}\mathbf{u}_r^{\mathrm{T}}/\|\mathbf{x}\| = \mathbf{x}\mathbf{u}_r^{\mathrm{T}}/\|\mathbf{x}\| = E^{\mathrm{T}}.$$

Note that, so far, we have not specified any structures for $\mathcal{S}$. So, it is not surprising that we cannot improve on κ. In the following sections, we will show that for some structures $\kappa^{\mathcal{S}}$ is smaller than κ, since φ can be smaller than $\|A^{+}\|\,\|\mathbf{x}\|$. We introduce a parameter in the relation between $\kappa^{\mathcal{S}}$ and φ. As we know, for any two vectors $\mathbf{u}$ and $\mathbf{v}$,

$$(\|\mathbf{u}\| + \|\mathbf{v}\|)/\sqrt{2} \leq \sqrt{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2} \leq \max(\|\mathbf{u}+\mathbf{v}\|,\ \|\mathbf{u}-\mathbf{v}\|) \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Introducing a parameter c, $1/\sqrt{2} \le c \le 1$, we have

$$\max(\|\mathbf{u}+\mathbf{v}\|,\ \|\mathbf{u}-\mathbf{v}\|) = c\,(\|\mathbf{u}\| + \|\mathbf{v}\|).$$

Applying the above equation to (2.4), we get

$$\|\Delta\mathbf{x}\| \approx c\,(\|A^+\Delta A\mathbf{x}\| + \|A^+\Delta\mathbf{b}\|),$$

for $1/\sqrt{2} \le c \le 1$, since we are free in choosing the sign of $\Delta\mathbf{b}$. Consequently, in (2.1),

$$\frac{\|\Delta\mathbf{x}\|}{\epsilon\|\mathbf{x}\|} \approx c\left(\frac{\|A^+\Delta A\mathbf{x}\|}{\epsilon\|\mathbf{x}\|} + \frac{\|A^+\Delta\mathbf{b}\|}{\epsilon\|\mathbf{x}\|}\right) = c\left(\frac{\|A\|\,\|A^+E\mathbf{x}\|}{\|\mathbf{x}\|} + \frac{\|A^+\Delta\mathbf{b}\|}{\epsilon\|\mathbf{x}\|}\right).$$

To eliminate ϵ in the above expression, on the one hand, obviously, we have

$$\frac{\|A^+\Delta\mathbf{b}\|}{\epsilon\|\mathbf{x}\|} \le \frac{\epsilon\|A^+\|\,\|\mathbf{b}\|}{\epsilon\|\mathbf{x}\|} = \|A^+\|\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}.$$

On the other hand, if we set $\Delta\mathbf{b} = \epsilon\|\mathbf{b}\|\mathbf{y}$, where $\mathbf{y}$ is a unit norm vector such that $\|A^+\mathbf{y}\| = \|A^+\|$, then for this particular $\Delta\mathbf{b}$ we have

$$\frac{\|A^+\Delta\mathbf{b}\|}{\epsilon\|\mathbf{x}\|} = \frac{\epsilon\|A^+\mathbf{y}\|\,\|\mathbf{b}\|}{\epsilon\|\mathbf{x}\|} = \|A^+\|\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}.$$

Using the definitions (2.1) and (2.8) of $\kappa^{\mathcal{S}}$ and φ respectively, we have the following theorem.

THEOREM 2.1. *Using the notations given above, the structured condition number*

$$\kappa^{\mathcal{S}} = c\left(\varphi\frac{\|A\|}{\|\mathbf{x}\|} + \|A^+\|\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}\right), \tag{2.11}$$

where $1/\sqrt{2} \le c \le 1$. *For no perturbation on the right hand side we have*

$$\kappa_0^{\mathcal{S}} = \varphi\,\frac{\|A\|}{\|\mathbf{x}\|}.$$

So, we focus our analysis of structured condition numbers to the analysis of φ. In the following sections, we discuss the condition numbers for specific structured matrices based on the general results presented in this section.

3 Symmetric, persymmetric and skewsymmetric.

In this section, we discuss the following structures: symmetric, persymmetric, and skewsymmetric. We show that there is no improvement on the condition number. In other words, $\kappa^{\mathcal{S}} = \kappa$ when $\mathcal{S} = \{$symmetric, persymmetric, and skewsymmetric$\}$. We first restate a lemma from [8].

LEMMA 3.1 ([8]). *Let $\mathcal{S} = \{$symmetric, persymmetric$\}$ and $\mathbf{x}, \mathbf{y} \in R^n$ such that $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$, then there exits an n-by-n matrix $A \in \mathcal{S}$ with*

$$\mathbf{y} = A\mathbf{x} \quad \textit{and} \quad \|A\| = 1. \tag{3.1}$$

If, in addition, $\mathbf{y}^{\mathrm{T}}\mathbf{x} = 0$, then there exists a skewsymmetric A satisfying (3.1).

If A is persymmetric, then JA is symmetric, where J is the permutation matrix

$$J := \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}. \tag{3.2}$$

Thus, the following lemma implies that symmetric and persymmetric least squares problems have the same value of φ, thus same condition number.

LEMMA 3.2. *The least squares problems*

$$\|A\mathbf{x} - \mathbf{b}\| = \min \quad \textit{and} \quad \|JA\mathbf{x} - J\mathbf{b}\| = \min$$

have the same value of φ.

PROOF. Obviously, $\|E\| = \|JE\|$ and if E and A have the same structure, then JE and JA have the same structure. It can be verified that $(JA)^+ = A^+J$. We then get $\|A^+E\mathbf{x}\| = \|(JA)^+(JE)\mathbf{x}\|$,

$$AA^+E = E \quad \Leftrightarrow \quad (JA)(JA)^+(JE) = JE,$$

and

$$A^+AE^{\mathrm{T}} = E^{\mathrm{T}} \quad \Leftrightarrow \quad (JA)^+(JA)(JE)^{\mathrm{T}} = (JE)^{\mathrm{T}}.$$

This completes the proof. $\square$

Then we present our result.

THEOREM 3.3. *For the structure $\mathcal{S} = \{$symmetric, persymmetric, skewsymmetric$\}$,*

$$\kappa^{\mathcal{S}} = \kappa.$$

PROOF. From (2.10), we only need to show that $\varphi = \|A^+\|\,\|\mathbf{x}\|$ when $\mathcal{S} = \{$symmetric, persymmetric, skewsymmetric$\}$. Since, from Lemma 3.2, symmetric and persymmetric have the same condition number, we consider two cases: symmetric and skewsymmetric.

Case 1 (symmetric). Let $r = \mathrm{rank}(A)$ and

$$A = V \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} V^{\mathrm{T}}$$

be the SVD of A. We partition $V = [V_1 \ V_2]$, where V_1 consists of the first r columns of V. Then

$$A^+ = [V_1 \ V_2] \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{bmatrix}$$

is the SVD of A^+ and it has a short form $A^+ = V_1\Sigma_r^{-1}V_1^{\mathrm{T}}$. Thus

$$\|V_1^{\mathrm{T}}\mathbf{x}\| = \|V_1^{\mathrm{T}}A^+\mathbf{b}\| = \|\Sigma_r^{-1}V_1^{\mathrm{T}}\mathbf{b}\| = \|V_1\Sigma_r^{-1}V_1^{\mathrm{T}}\mathbf{b}\| = \|\mathbf{x}\|.$$

Applying Lemma 3.1 to the vectors $V_1^{\mathrm{T}}\mathbf{x}/\|\mathbf{x}\|$ and $\mathbf{e}_r$, the rth column of the identity matrix, we have a symmetric matrix $\widehat{E}$ such that

$$\widehat{E}V_1^{\mathrm{T}}\mathbf{x}/\|\mathbf{x}\| = \mathbf{e}_r \quad \text{and} \quad \|\widehat{E}\| = 1.$$

Let $E = V_1\widehat{E}V_1^{\mathrm{T}}$, then E is symmetric and

$$\|A^+E\mathbf{x}\| = \|A^+V_1\mathbf{e}_r\|\,\|\mathbf{x}\| = \|V_1\Sigma_r^{-1}\mathbf{e}_r\|\,\|\mathbf{x}\| = \|A^+\|\,\|\mathbf{x}\|.$$

Moreover, $AA^+E = (V_1V_1^{\mathrm{T}})(V_1\widehat{E}V_1^{\mathrm{T}}) = E$ and $A^+AE^{\mathrm{T}} = E^{\mathrm{T}}$ since $A^+A = AA^+ = V_1V_1^{\mathrm{T}}$ and E is symmetric. This means that we have constructed a symmetric E in (2.8) so that $\varphi = \|A^+\|\,\|\mathbf{x}\|$.

Case 2 (skewsymmetric). A matrix is skewsymmetric if $A = -A^{\mathrm{T}}$. It can be verified that $(A^{\mathrm{T}})^+ = (A^+)^{\mathrm{T}}$. So, $A = -A^{\mathrm{T}}$ implies that

$$A^+ = -(A^+)^{\mathrm{T}}, \tag{3.3}$$

in other words, A^+ is also skewsymmetric. Consequently,

$$AA^+ = -A(A^+)^{\mathrm{T}} = (A^+A)^{\mathrm{T}} = A^+A. \tag{3.4}$$

We know that the eigenvalues of a skewsymmetric matrix are either conjugate pure imaginary or zero. Also, all the singular values of a skewsymmetric matrix are of even multiplicity. Let

$$A = U\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}V^{\mathrm{T}}$$

be the SVD of A, then the SVD of A^+ is

$$A^+ = V\begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix}U^{\mathrm{T}}.$$

Partitioning $U = [\mathbf{u}_1\ \mathbf{u}_2 \cdots \mathbf{u}_n]$, we assume that $\mathbf{u}_1$ and $\mathbf{u}_2$ are the singular vectors corresponding to the largest singular value of A^+. Thus, $\|A^+\mathbf{u}_1\| = \|A^+\mathbf{u}_2\| = \|A^+\|$. Set

$$\mathbf{y} := \eta_1\mathbf{u}_1 + \eta_2\mathbf{u}_2.$$

If

$$\eta_1^2 + \eta_2^2 = 1, \tag{3.5}$$

then $\|\mathbf{y}\| = 1$ and $\|A^+\mathbf{y}\| = \|A^+\|$. Next, we find a skewsymmetric matrix E such that $E\mathbf{x}/\|\mathbf{x}\| = \mathbf{y}$. It then follows that $\|A^+E\mathbf{x}\| = \|A^+\mathbf{y}\|\,\|\mathbf{x}\| = \|A^+\|\,\|\mathbf{x}\|$.

Denote the vector

$$\mathbf{a} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} := \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} V^{\mathrm{T}}\mathbf{b},$$

then, from (3.3), we get

$$\mathbf{x} = A^{+}\mathbf{b} = -(A^{+})^{\mathrm{T}}\mathbf{b} = -U\mathbf{a}.$$

It then follows that

$$\mathbf{x}^{\mathrm{T}}\mathbf{y} = -\mathbf{a}^{\mathrm{T}}U^{\mathrm{T}}\mathbf{y} = -\alpha_1\eta_1 - \alpha_2\eta_2.$$

Choosing η_1 and η_2 such that

$$\alpha_1\eta_1 + \alpha_2\eta_2 = 0, \tag{3.6}$$

we have $(\mathbf{x}/\|\mathbf{x}\|)^{\mathrm{T}}\mathbf{y} = 0$. From Lemma 3.1, there exists a skewsymmetric matrix $\widehat{E}$, such that $\|\widehat{E}\| = 1$ and $\widehat{E}\mathbf{x}/\|\mathbf{x}\| = \mathbf{y}$ provided that η_1 and η_2 satisfy (3.5) and (3.6). Note that $\mathbf{x} = A^{+}\mathbf{b} = A^{+}A\mathbf{x}$. So, $\widehat{E}A^{+}A\mathbf{x}/\|\mathbf{x}\| = \mathbf{y}$. Define $E := A^{+}A\widehat{E}A^{+}A$, it can be verified that E is skewsymmetric and $\|E\| \le \|\widehat{E}\| = 1$, and also $E\mathbf{x}/\|\mathbf{x}\| = A^{+}A\mathbf{y}$. Thus, from (3.4), we have

$$\|A^{+}E\mathbf{x}\| = \|A^{+}A^{+}A\mathbf{y}\|\,\|\mathbf{x}\| = \|A^{+}\mathbf{y}\|\,\|\mathbf{x}\| = \|A^{+}\|\,\|\mathbf{x}\|.$$

Finally, we show that E satisfies the conditions (2.9):

$$AA^{+}E = AA^{+}\widehat{E}A^{+}A = E,$$

which implies

$$A^{+}AE^{\mathrm{T}} = -AA^{+}E = -E = E^{\mathrm{T}},$$

since E is skewsymmetric. $\square$

4 Exploiting structures.

It is not surprising that we are unable to improve the condition numbers for symmetric, persymmetric, and skewsymmetric structures. Those matrices have $O(n^2)$ free entries, same magnitude order as an n-by-n general matrix. However, matrices like Toeplitz and Hankel have $O(n)$ free entries. In this section, we consider these structures. First, we group the structures according to their condition numbers. A matrix T is symmetric Toeplitz if and only if JT (or TJ) is persymmetric Hankel, where J is the permutation matrix in (3.2). Thus, from Lemma 3.2, symmetric Toeplitz and persymmetric Hankel least squares problems have the same condition number. Also, Hankel and (general) Toeplitz least squares problems have the same condition number since a matrix T is Toeplitz

Table 4.1: The number of independent parameters for the structured matrices.

structure	symmetric	skewsymmetric	circulant	symmetric Toeplitz	Hankel
k	$(n^2+n)/2$	$(n^2-n)/2$	n	n	$2n-1$

if and only if JT (or TJ) is Hankel. In this section, we present a general result on the condition number for the structures: circulant, symmetric Toeplitz, and Hankel.

To exploit structures, we extract the independent entries of a matrix into a vector $\mathbf{p}$ and decompose the matrix into a product of a matrix Φ representing the structure of the matrix and the parameter vector $\mathbf{p}$. Specifically, we denote $\mathrm{vec}(A)$ as the vector of stacked columns of A, then the vector $\mathbf{p}$ is the subvector of $\mathrm{vec}(A)$ consisting of the independent entries of A. The matrix Φ only contains the information about the structure of A such that

$$\mathrm{vec}(A) = \Phi\mathbf{p}. \tag{4.1}$$

For example, if A is an n-by-n Toeplitz matrix and $n = 3$:

$$A = \begin{bmatrix} t_0 & t_{-1} & t_{-2} \\ t_1 & t_0 & t_{-1} \\ t_2 & t_1 & t_0 \end{bmatrix},$$

then A has $2n-1$ independent entries and $\mathbf{p}$ is a $(2n-1)$-vector:

$$\mathbf{p} = \begin{bmatrix} t_0 \\ t_1 \\ t_2 \\ t_{-1} \\ t_{-2} \end{bmatrix},$$

which is the subvector of $\mathrm{vec}(A)$ consisting of the first $2n-1$ independent entries. The n^2-by-$(2n-1)$ matrix Φ, which is sparse and has entries 0 or 1, satisfies (4.1). Table 4.1 shows the number k of independent parameters, or the dimensions of the parameter vector $\mathbf{p}$, for the matrices under investigation.

Now, a structured perturbation ΔA on a structured matrix A is determined by a perturbation $\Delta\mathbf{p}$ on the parameter vector $\mathbf{p}$. Specifically, if $\mathrm{vec}(A) = \Phi\mathbf{p}$, then $\mathrm{vec}(\Delta A) = \Phi(\Delta\mathbf{p})$. To exploit structures, we replace the matrix E in (2.8) with a k-vector $\mathbf{t}$ defined by $\mathbf{t} = \Delta\mathbf{p}/(\epsilon\,\|A\|)$. Recalling that $E = \Delta A/(\epsilon\,\|A\|)$, we have $\mathrm{vec}(E) = \Phi\mathbf{t}$. The following lemma [8] establishes the relations between $\|E\|$ (or $\|A\|$) and $\|\mathbf{t}\|$ (or $\|\mathbf{p}\|$) for the structures under investigation.

LEMMA 4.1 ([8]). *Let* $\mathrm{vec}(E) = \Phi\mathbf{t}$ *be the decomposition as described above, then*

$$\alpha\,\|E\| \le \|\mathbf{t}\| \le \beta\,\|E\|, \tag{4.2}$$

Table 4.2: The constants α and β in (4.2) for the structures: circulant, symmetric Toeplitz, and Hankel.

structures	circulant	symmetric Toeplitz	Hankel
α	$1/\sqrt{n}$	$1/\sqrt{2n-2}$	$1/\sqrt{n}$
β	1	1	$\sqrt{2}$

where the constants α and β are given by Table 4.2. The upper bounds for all the structures are sharp. The lower bound for circulant structure is sharp and the lower bounds for symmetric Toeplitz and Hankel are sharp up to a factor of $\sqrt{2}$.

It follows from the above lemma that

$$\begin{aligned} &\{E \in \mathcal{S} : \ \mathrm{vec}(E) = \Phi\mathbf{t}, \ \|\mathbf{t}\| \le \alpha\} \\ &\quad \subseteq \{E \in \mathcal{S} : \ \|E\| \le 1\} \\ &\quad \subseteq \{E \in \mathcal{S} : \ \mathrm{vec}(E) = \Phi\mathbf{t}, \ \|\mathbf{t}\| \le \beta\}. \end{aligned} \tag{4.3}$$

So, we can replace the condition $\|E\| \le 1$ with a condition on $\|\mathbf{t}\|$. Further, in (2.8), we note that

$$E\mathbf{x} = (\mathbf{x}^{\mathrm{T}} \otimes I)\mathrm{vec}(E) = (\mathbf{x}^{\mathrm{T}} \otimes I)\Phi\mathbf{t},$$

where $\mathbf{x}^{\mathrm{T}} \otimes I$ is the Kronecker product. Defining an n-by-k matrix

$$\Psi := (\mathbf{x}^{\mathrm{T}} \otimes I)\Phi,$$

which is independent of the perturbation ΔA, from (4.3), we have $\varphi \le \sup\{\|A^{+}\Psi\mathbf{t}\| : \ \|\mathbf{t}\| \le \beta\} = \beta\,\|A^{+}\Psi\|$. Thus we get

$$\varphi = \gamma\,\|A^{+}\Psi\|, \quad \text{for} \quad 0 \le \gamma \le \beta.$$

Because of the parameter γ, we may expect a smaller φ for the structures such as circulant, symmetric Toeplitz, and Hankel. From (2.11), we have the following result on the condition number for circulant, symmetric Toeplitz, and Hankel.

THEOREM 4.2. *For $\mathcal{S} = \{$circulant, symmetric Toeplitz, Hankel$\}$,*

$$\kappa^{\mathcal{S}} = c\left(\gamma\,\|A^{+}\Psi\|\frac{\|A\|}{\|\mathbf{x}\|} + \|A^{+}\|\frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}\right),$$

where $1/\sqrt{2} \le c \le 1$ and $0 \le \gamma \le \beta$ for β given in Lemma 4.1. In the case of no perturbation on the right hand side,

$$\kappa_0^{\mathcal{S}} = \gamma\,\|A^{+}\Psi\|\frac{\|A\|}{\|\mathbf{x}\|}.$$

This implies the following ratios of the structured and unstructured condition numbers.

COROLLARY 4.3. *For $\mathcal{S} = \{$circulant, symmetric Toeplitz, Hankel$\}$,*

$$\frac{\kappa^{\mathcal{S}}}{\kappa} \geq (1/\sqrt{2}) \frac{\|A^+\| \, \|\mathbf{b}\|/\|\mathbf{x}\|}{\|A^+\| \, \|A\| + \|A^+\| \, \|A\mathbf{x}\|/\|\mathbf{x}\|}.$$

5 Circulant.

An n-by-n circulant matrix has the form:

$$C = \mathrm{circ}(c_0, c_1, ..., c_{n-1}) = \begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_1 \\ c_1 & \cdots & c_{n-1} & c_0 \end{bmatrix}. \tag{5.1}$$

Circulant matrices have a number of remarkable properties [3]. In particular, denote the circularly shift-up matrix

$$P = \begin{bmatrix} 0 & 1 & & 0 \\ \vdots & \ddots & \ddots & \\ 0 & \ddots & \ddots & 1 \\ 1 & 0 & \cdots & 0 \end{bmatrix},$$

then the circulant matrix C in (5.1) can be written as

$$C = \sum_{k=0}^{n-1} c_k P^k.$$

This polynomial representation implies that circulant matrices commute. Also, C can be diagonalized by Fourier transformation: $C = F^{-1}DF$, where F is the discrete Fourier transformation and D is diagonal. It then follows that the generalized inverse C^+ is also circulant and the product of circulant matrices is circulant. Particularly, if A is circulant, the AA^+ is circulant. So, if we choose $E = AA^+$ in the definition (2.8) of φ, then $\|A^+E\mathbf{x}\| = \|A^+\mathbf{x}\|$, which implies that $\varphi \geq \|A^+\mathbf{x}\|$. On the other hand, since both A^+ and E are circulant, they commute. Thus $\|A^+E\mathbf{x}\| = \|EA^+\mathbf{x}\| \leq \|A^+\mathbf{x}\|$. Therefore,

$$\varphi^{\mathrm{circ}} = \|A^+\mathbf{x}\|. \tag{5.2}$$

Using the above value of φ in (2.11), we have the following result of the condition number for the circulant least squares problem.

THEOREM 5.1. *For the circulant least squares problem, the condition number*

$$\kappa^{\mathrm{circ}} = c\left(\|A^+\mathbf{x}\| \frac{\|A\|}{\|\mathbf{x}\|} + \|A^+\| \frac{\|\mathbf{b}\|}{\|\mathbf{x}\|} \right),$$

where $1/\sqrt{2} \le c \le 1$. *For no perturbation on the right hand side,*

$$\kappa_0^{\rm circ} = \|A^+\mathbf{x}\| \frac{\|A\|}{\|\mathbf{x}\|}$$

and, from (1.4),

$$\frac{\kappa_0^{\rm circ}}{\kappa_0} = \frac{\|A^+\mathbf{x}\|}{\|A^+\|\,\|\mathbf{x}\|} \ge \frac{1}{\|A^+\|\,\|A\|}.$$

PROOF. The last inequality follows from $\|\mathbf{x}\| \le \|A\|\,\|A^+\mathbf{x}\|$. □

Before comparing the circulant condition number $\kappa^{\rm circ}$ and the general condition number κ in (1.3), we give two lemmas. The first is from [8]. The second gives a simple lower bound on $\kappa^{\mathcal{S}}$.

LEMMA 5.2 ([8]). *For any* n*-by-*n *matrix* A*,* n*-vector* $\mathbf{z}$*, and* n*-by-*n *circulant matrix* C*,*

$$\|AC\| = \|AC^{\rm H}\| \quad \text{and} \quad \|C\mathbf{z}\| = \|C^{\rm H}\mathbf{z}\|.$$

LEMMA 5.3. *For an* n*-by-*n *matrix* A *and structure* $\mathcal{S}$*, if*

$$\varphi \ge \omega\,\|(A^+)^{\rm T}\mathbf{x}\|, \tag{5.3}$$

for some nonnegative ω*, then*

$$\kappa^{\mathcal{S}} \ge \sqrt{2\omega\,\|A^+\|\,\|A\|}\,. \tag{5.4}$$

PROOF. Since $\mathbf{x}$ is the least norm least squares solution,

$$\|\mathbf{x}\|^2 = \mathbf{x}^{\rm T}\mathbf{x} = \mathbf{x}^{\rm T}A^+AA^+\mathbf{b} \le \|\mathbf{x}^{\rm T}A^+\|\,\|AA^+\|\,\|\mathbf{b}\| = \|(A^+)^{\rm T}\mathbf{x}\|\,\|\mathbf{b}\|.$$

From the condition (5.3) and the above equality, we get

$$\begin{aligned}\varphi\,\|A\| + \|A^+\|\,\|\mathbf{b}\| &\ge \omega\,\|(A^+)^{\rm T}\mathbf{x}\|\,\|A\| + \|A^+\|\,\|\mathbf{b}\| \\ &\ge 2\sqrt{\omega\,\|(A^+)^{\rm T}\mathbf{x}\|\,\|A\|\,\|A^+\|\,\|\mathbf{b}\|} \\ &\ge 2\,\|\mathbf{x}\|\sqrt{\omega\,\|A^+\|\,\|A\|}\,.\end{aligned}$$

Then, (5.4) follows from (2.11) in Theorem 2.1. □

Finally, we have a simple ratio between $\kappa^{\rm circ}$ and κ.

THEOREM 5.4. *For the condition numbers for the circulant and general least squares problems, we have*

$$\frac{\kappa^{\rm circ}}{\kappa} \ge \frac{1}{\sqrt{2\|A^+\|\,\|A\|}}\,. \tag{5.5}$$

PROOF. Using $\varphi^{\rm circ}$ in (5.2) and Lemma 5.2, we have

$$\varphi^{\rm circ} = \|A^+\mathbf{x}\| = \|(A^+)^{\rm T}\mathbf{x}\|.$$

It then follows from Lemma 5.3 that

$$\kappa^{\text{circ}} \geq \sqrt{2\|A^+\|\,\|A\|}\,,$$

which, from (1.3), implies (5.5). ◻

This theorem gives a lower bound for the ratio $\kappa^{\text{circ}}/\kappa$ and shows that we may expect a smaller κ^{circ}. It is shown by an example in [8] that, for nonsingular A, the lower bound (5.5) for the ratio $\kappa^{\text{circ}}/\kappa$ is sharp.

6 Symmetric Toeplitz.

First, we assume n is even and $n = 2m$. The case when n is odd will be considered later in this section. If A is symmetric Toeplitz, we have the partition

$$A = \begin{bmatrix} T & U \\ U^{\mathrm{T}} & T \end{bmatrix}, \tag{6.1}$$

where T and U are matrices of order m and T is symmetric. Accordingly, we partition

$$J = \begin{bmatrix} 0 & J_{1/2} \\ J_{1/2} & 0 \end{bmatrix}, \tag{6.2}$$

where $J_{1/2}$ is m-by-m. Using the above partitions (6.1) and (6.2) and $JAJ = A$, we can verify that $T = J_{1/2}TJ_{1/2}$ and $U^{\mathrm{T}} = J_{1/2}UJ_{1/2}$.

Second, we assume that the solution $\mathbf{x}$ satisfies $\mathbf{x} = \pm J\mathbf{x}$. This assumption will also be removed.

LEMMA 6.1. *For a matrix B satisfying $B = JBJ$ and a vector $\mathbf{c}$ satisfying $\mathbf{c} = \pm J\mathbf{c}$, the vector $\mathbf{d} = B\mathbf{c}$ also satisfies $\mathbf{d} = \pm J\mathbf{d}$. For the generalized inverse B^+, we also have $B^+\mathbf{c} = \pm J(B^+\mathbf{c})$.*

PROOF. This follows from

$$\pm J\mathbf{d} = \pm JB\mathbf{c} = (JBJ)(\pm J\mathbf{c}) = B\mathbf{c} = \mathbf{d}$$

and note that $B^+ = (JBJ)^+ = JB^+J$. ◻

Clearly, if $\mathbf{c} = \pm J\mathbf{c}$, then $\mathbf{c}$ has the form:

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_1 \\ \pm J_{1/2}\mathbf{c}_1 \end{bmatrix},$$

where $\mathbf{c}_1$ is an m-vector, the top half of $\mathbf{c}$.

Now, we consider $A^+E\mathbf{x}$ in (2.8). Denote $\mathbf{y} := E\mathbf{x}$. Since E is symmetric Toeplitz, it can be verified that $JEJ = E$. From the assumption $\mathbf{x} = \pm J\mathbf{x}$ and Lemma 6.1, the vector $\mathbf{y}$ has the form

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \pm J_{1/2}\mathbf{y}_1 \end{bmatrix}.$$

Denoting $\mathbf{z} := A^+E\mathbf{x} = A^+\mathbf{y}$, again from Lemma 6.1, the vector $\mathbf{z}$ also has the form

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ \pm J_{1/2}\mathbf{z}_1 \end{bmatrix}.$$

If A^+ is partitioned as

$$A^+ = \begin{bmatrix} T_1 & U_1 \\ U_1^{\mathrm{T}} & T_2 \end{bmatrix}, \tag{6.3}$$

then $\mathbf{z} = A^+\mathbf{y}$ implies that

$$\mathbf{z}_1 = (T_1 \pm U_1 J_{1/2})\mathbf{y}_1. \tag{6.4}$$

Using the partition (6.3) and $A^+ = JA^+J$, we can verify that $T_2 = J_{1/2}T_1J_{1/2}$ and $U_1^{\mathrm{T}} = J_{1/2}U_1J_{1/2}$.

In the following, we establish a relation between $T_1 \pm U_1J_{1/2}$ and $T \pm UJ_{1/2}$. Specifically, we show that $T_1 \pm U_1J_{1/2}$ is the Moore–Penrose inverse of $T \pm UJ_{1/2}$. We first present the following two lemmas.

LEMMA 6.2. *If* $\mathbf{u} \in \mathrm{Range}(A)$ *and has the form*

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \pm J_{1/2}\mathbf{u}_1 \end{bmatrix},$$

then

$$\mathbf{u}_1 = (T \pm UJ_{1/2})(T_1 \pm U_1J_{1/2})\mathbf{u}_1 = (T_1 \pm U_1J_{1/2})(T \pm UJ_{1/2})\mathbf{u}_1. \tag{6.5}$$

PROOF. From Lemma 6.1, if $\mathbf{u} = \pm J\mathbf{u}$, then both $\mathbf{v} := A\mathbf{u}$ and $\mathbf{w} := A^+\mathbf{u}$ have the form:

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \pm J_{1/2}\mathbf{v}_1 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \pm J_{1/2}\mathbf{w}_1 \end{bmatrix}.$$

From the partitions (6.1) and (6.3), $\mathbf{v}_1 = (T \pm UJ_{1/2})\mathbf{u}_1$ and $\mathbf{w}_1 = (T_1 \pm U_1J_{1/2})\mathbf{u}_1$. When $\mathbf{u} \in \mathrm{Range}(A)$,

$$\mathbf{u} = AA^+\mathbf{u} = A^+A\mathbf{u},$$

since A is symmetric. We then have

$$\mathbf{u} = AA^+\mathbf{u} = A\mathbf{w} \quad \text{and} \quad \mathbf{u} = A^+A\mathbf{u} = A^+\mathbf{v}.$$

From the partitions (6.1) and (6.3), we have $\mathbf{u}_1 = (T \pm UJ_{1/2})\mathbf{w}_1$ and $\mathbf{u}_1 = (T_1 \pm U_1J_{1/2})\mathbf{v}_1$, and then (6.5) follows. $\square$

LEMMA 6.3. *If* $\mathbf{u}_1 \in \mathrm{Range}(T \pm UJ_{1/2})$ *or* $\mathbf{u}_1 \in \mathrm{Range}(T_1 \pm U_1J_{1/2})$ *and*

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_1 \\ \pm J_{1/2}\mathbf{u}_1 \end{bmatrix},$$

then $\mathbf{u} \in \mathrm{Range}(A)$.

PROOF. When $\mathbf{u}_1 \in \mathrm{Range}(T \pm UJ_{1/2})$, let $\mathbf{u}_1 = (T \pm UJ_{1/2})\mathbf{t}$ for some $\mathbf{t} \in R^m$, then

$$\mathbf{u} = \begin{bmatrix} (T \pm UJ_{1/2})\mathbf{t} \\ \pm J_{1/2}(T \pm UJ_{1/2})\mathbf{t} \end{bmatrix} = \begin{bmatrix} (T \pm UJ_{1/2})\mathbf{t} \\ (U^{\mathrm{T}} \pm TJ_{1/2})\mathbf{t} \end{bmatrix} = A \begin{bmatrix} \mathbf{t} \\ \pm J_{1/2}\mathbf{t} \end{bmatrix},$$

noting that $U^{\mathrm{T}} = J_{1/2}UJ_{1/2}$, $T = J_{1/2}TJ_{1/2}$, and $J_{1/2}J_{1/2} = I$. Thus, $\mathbf{u} \in \mathrm{Range}(A)$. Similarly, when $\mathbf{u}_1 \in \mathrm{Range}(T_1 \pm U_1J_{1/2})$, then we can show that $\mathbf{u} \in \mathrm{Range}(A^+)$, which implies that $\mathbf{u} \in \mathrm{Range}(A)$ since A is symmetric. □

It follows from Lemma 6.2 and Lemma 6.3 that if $\mathbf{u}_1 \in \mathrm{Range}(T \pm UJ_{1/2})$ or $\mathbf{u}_1 \in \mathrm{Range}(T_1 \pm U_1J_{1/2})$, then the equations in (6.5) hold. Thus, for any $\mathbf{t} \in R^m$,

$$(T \pm UJ_{1/2})(T_1 \pm U_1J_{1/2})(T \pm UJ_{1/2})\mathbf{t} = (T \pm UJ_{1/2})\mathbf{t}$$

and

$$(T_1 \pm U_1J_{1/2})(T \pm UJ_{1/2})(T_1 \pm U_1J_{1/2})\mathbf{t} = (T_1 \pm U_1J_{1/2})\mathbf{t},$$

which implies that

$$(T \pm UJ_{1/2})(T_1 \pm U_1J_{1/2})(T \pm UJ_{1/2}) = T \pm UJ_{1/2}$$

and

$$(T_1 \pm U_1J_{1/2})(T \pm UJ_{1/2})(T_1 \pm U_1J_{1/2}) = T_1 \pm U_1J_{1/2}.$$

In other words, $T_1 \pm U_1J_{1/2}$ is a $\{1, 2\}$-inverse of $T \pm UJ_{1/2}$ [2].

We then show that $T \pm UJ_{1/2}$ and $T_1 \pm U_1J_{1/2}$ commute, thus $T_1 \pm U_1J_{1/2}$ is the group inverse of $T \pm UJ_{1/2}$ [2]. First, we show that $\mathrm{Range}(T \pm UJ_{1/2}) \cap \mathrm{Null}(T \pm UJ_{1/2}) = \{0\}$. Indeed, if $\mathbf{t} \in \mathrm{Range}(T \pm UJ_{1/2})$, then from Lemmas 6.2 and 6.3,

$$\mathbf{t} = (T_1 \pm U_1J_{1/2})(T \pm UJ_{1/2})\mathbf{t}.$$

So, if $\mathbf{t}$ is also in $\mathrm{Null}(T \pm UJ_{1/2})$, then $\mathbf{t} = 0$. Then, we know that $\mathrm{Range}(T \pm UJ_{1/2})$ and $\mathrm{Null}(T \pm UJ_{1/2})$ are complementary subspaces of R^m, because $\dim(\mathrm{Range}(T \pm UJ_{1/2})) + \dim(\mathrm{Null}(T \pm UJ_{1/2})) = m$. Finally, we can prove that $T \pm UJ_{1/2}$ and $T_1 \pm U_1J_{1/2}$ commute. From Lemmas 6.2 and 6.3, for any $\mathbf{u}_1 \in \mathrm{Range}(T \pm UJ_{1/2})$ or $\mathbf{u}_1 \in \mathrm{Range}(T_1 \pm U_1J_{1/2})$, (6.5) holds. Thus, $\mathrm{Range}(T \pm UJ_{1/2}) = \mathrm{Range}(T_1 \pm U_1J_{1/2})$. Consequently, we also have $\mathrm{Null}(T \pm UJ_{1/2}) = \mathrm{Null}(T_1 \pm U_1J_{1/2})$, since $\mathrm{Range}(T \pm UJ_{1/2}) \oplus \mathrm{Null}(T \pm UJ_{1/2}) = \mathrm{Range}(T_1 \pm U_1J_{1/2}) \oplus \mathrm{Null}(T_1 \pm U_1J_{1/2}) = R^m$. Now, for any vector $\mathbf{x} \in R^m$, we can write it as $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$, where $\mathbf{x}_1 \in \mathrm{Range}(T \pm UJ_{1/2}) = \mathrm{Range}(T_1 \pm U_1 J_{1/2})$ and $\mathbf{x}_2 \in \mathrm{Null}(T \pm UJ_{1/2}) = \mathrm{Null}(T_1 \pm U_1 J_{1/2})$. Thus,

$$\begin{aligned}(T \pm UJ_{1/2})(T_1 \pm U_1 J_{1/2})\mathbf{x} &= (T \pm UJ_{1/2})(T_1 \pm U_1 J_{1/2})(\mathbf{x}_1 + \mathbf{x}_2) \\ &= (T \pm UJ_{1/2})(T_1 \pm U_1 J_{1/2})\mathbf{x}_1 \\ &= (T_1 \pm U_1 J_{1/2})(T \pm UJ_{1/2})\mathbf{x}_1 \\ &= (T_1 \pm U_1 J_{1/2})(T \pm UJ_{1/2})\mathbf{x},\end{aligned}$$

which implies that $T \pm UJ_{1/2}$ and $T_1 \pm U_1 J_{1/2}$ commute.

In this case, since T is symmetric and $U^{\mathrm{T}} = J_{1/2} U J_{1/2}$, $T \pm UJ_{1/2}$ is symmetric, which means the group inverse of $T \pm UJ_{1/2}$ is its Moore–Penrose inverse [2]. Thus, from (6.4), we have

$$\mathbf{z}_1 = (T \pm UJ_{1/2})^+ \mathbf{y}_1.$$

Now, we have

$$\begin{aligned}\|A^+ E\mathbf{x}\| &= \left\| \begin{bmatrix} \mathbf{z}_1 \\ \pm J_{1/2}\mathbf{z}_1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} (T \pm UJ_{1/2})^+ \mathbf{y}_1 \\ \pm J_{1/2}(T \pm UJ_{1/2})^+ \mathbf{y}_1 \end{bmatrix} \right\| \\ &\le \|(T \pm UJ_{1/2})^+\| \left\| \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_1 \end{bmatrix} \right\|.\end{aligned}$$

Finally,

$$\varphi \le \|(T \pm UJ_{1/2})^+\| \, \|\mathbf{x}\|,$$

since

$$\left\| \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_1 \end{bmatrix} \right\| = \|\mathbf{y}\| = \|E\mathbf{x}\| \le \|\mathbf{x}\|.$$

Now, we consider the case when n is odd, $n = 2m - 1$ for some integer m. Partition

$$A = \begin{bmatrix} T & U \\ U^{\mathrm{T}} & \widehat{T} \end{bmatrix} \quad \text{and} \quad A^+ = \begin{bmatrix} T_1 & U_1 \\ U_1^{\mathrm{T}} & T_2 \end{bmatrix}, \tag{6.6}$$

where T and T_1 are symmetric and of order $m = \lceil n/2 \rceil$, then, in this case, $\widehat{T}$ and T_2 are symmetric and of order $m - 1$ and U and U_1 are m-by-$(m-1)$. Defining an $(m-1)$-by-m matrix

$$\widehat{J}_{1/2} = \begin{bmatrix} 0 & & 1 & 0 \\ & \iddots & & \vdots \\ 1 & & 0 & 0 \end{bmatrix}, \tag{6.7}$$

we partition J as

$$J = \begin{bmatrix} D & \widehat{J}_{1/2}^{\mathrm{T}} \\ \widehat{J}_{1/2} & 0 \end{bmatrix},$$

where

$$D = \begin{bmatrix} 0 & & & 0 \\ & \ddots & & \\ & & 0 & \\ 0 & & & 1 \end{bmatrix}.$$

From the structure of $\widehat{J}_{1/2}$, we get

$$\widehat{J}_{1/2}\widehat{J}_{1/2}^{\mathrm{T}} = I_{m-1}, \quad \text{and} \quad \widehat{J}_{1/2}^{\mathrm{T}}\widehat{J}_{1/2} = I_m - D.$$

From the partitions in (6.6) and $JAJ = A$ and $JA^+J = A^+$, we can verify that

$$\widehat{T} = \widehat{J}_{1/2}T\widehat{J}_{1/2}^{\mathrm{T}}, \quad U^{\mathrm{T}} = \widehat{J}_{1/2}TD + \widehat{J}_{1/2}U\widehat{J}_{1/2}$$

and

$$T_2 = \widehat{J}_{1/2}T_1\widehat{J}_{1/2}^{\mathrm{T}}, \quad U_1^{\mathrm{T}} = \widehat{J}_{1/2}T_1D + \widehat{J}_{1/2}U_1\widehat{J}_{1/2}.$$

Also, in this case of odd n, if a vector $\mathbf{c} = \pm J\mathbf{c}$, then $\mathbf{c}$ has the form:

$$\mathbf{c} = \begin{bmatrix} \mathbf{c}_1 \\ \pm\widehat{J}_{1/2}\mathbf{c}_1 \end{bmatrix},$$

where $\mathbf{c}_1$ is an m-vector. Then (6.4) becomes

$$\mathbf{z}_1 = \big(T_1 \pm U_1\widehat{J}_{1/2}\big)\mathbf{y}_1.$$

For Lemma 6.2, it is easy to see that

$$\mathbf{u}_1 = \big(T \pm U\widehat{J}_{1/2}\big)\big(T_1 \pm U_1\widehat{J}_{1/2}\big)\mathbf{u}_1 = \big(T_1 \pm U_1\widehat{J}_{1/2}\big)\big(T \pm U\widehat{J}_{1/2}\big)\mathbf{u}_1.$$

As for Lemma 6.3, we can show that if $\mathbf{u}_1 \in \mathrm{Range}(T \pm U\widehat{J}_{1/2})$ or $\mathbf{u}_1 \in \mathrm{Range}(T_1 \pm U_1\widehat{J}_{1/2})$ then $\mathbf{u} \in \mathrm{Range}(A)$. Specifically,

$$\begin{aligned} \mathbf{u} &= \begin{bmatrix} \big(T \pm U\widehat{J}_{1/2}\big)\mathbf{t} \\ \big(\widehat{J}_{1/2}U\widehat{J}_{1/2} \pm \widehat{J}_{1/2}T\big)\mathbf{t} \end{bmatrix} \\ &= \begin{bmatrix} \big(T \pm U\widehat{J}_{1/2}\big)\mathbf{t} \\ \big(U^{\mathrm{T}} \pm \widehat{T}\widehat{J}_{1/2}\big)\mathbf{t} \end{bmatrix} \\ &= A\begin{bmatrix} \mathbf{t} \\ \pm\widehat{J}_{1/2}\mathbf{t} \end{bmatrix}, \end{aligned}$$

noting that $U^{\mathrm{T}} = \widehat{J}_{1/2}TD + \widehat{J}_{1/2}U\widehat{J}_{1/2}$, $\widehat{J}_{1/2}^{\mathrm{T}}\widehat{J}_{1/2} = I_m - D$, and $\widehat{T} = \widehat{J}_{1/2}T\widehat{J}_{1/2}^{\mathrm{T}}$. Then, analogous to the case of even n, $T_1 \pm U_1\widehat{J}_{1/2}$ is the group inverse of $T \pm U\widehat{J}_{1/2}$ and

$$\mathbf{z}_1 = \big(T \pm U\widehat{J}_{1/2}\big)^{\#}\mathbf{y}_1,$$

where $A^{\#}$ denotes the group inverse of A. Consequently,

$$\varphi \leq \|(T \pm U\widehat{J}_{1/2})^{\#}\| \, \|\mathbf{x}\|.$$

Putting things together, redefine

$$J_{1/2} = \begin{cases} J_{1/2} \text{ in (6.2)} & \text{when } n \text{ is even} \\ \widehat{J}_{1/2} \text{ in (6.7)} & \text{when } n \text{ is odd} \end{cases}$$

then we can unify the two cases.

THEOREM 6.4. *For $\mathcal{S} = \{$symmetric Toeplitz$\}$, assume A is partitioned as in (6.6), where $m = \lceil n/2 \rceil$, and the solution $\mathbf{x}$ satisfies $\mathbf{x} = J\mathbf{x}$, then the condition number*

$$\kappa^{\mathcal{S}} \leq \left\|(T + UJ_{1/2})^{\#}\right\| \, \|A\| + \|A^{+}\| \frac{\|\mathbf{b}\|}{\|\mathbf{x}\|}.$$

For no perturbation on the right hand side,

$$\frac{\kappa_0^{\mathcal{S}}}{\kappa_0} \leq \frac{\left\|(T + UJ_{1/2})^{\#}\right\|}{\|A^{+}\|}.$$

Note that when n is even, $T + UJ_{1/2}$ is symmetric and $(T + UJ_{1/2})^{\#} = (T + UJ_{1/2})^{+}$.

It arises in signal processing that the first column of a symmetric Toeplitz matrix is determined by

$$a_k = \sum_{j=1}^{M} t_j^2 \cos((k-1)\omega_j).$$

In this example, we set $n = 6$, $M = 2$, $\omega_1 = 0.63822\pi$, $\omega_2 = 0.0022817\pi$, $t_1 = 0.20766$, and $t_2 = 0.16112$. Then the first column of the symmetric Toeplitz matrix is given by

$$\begin{bmatrix} 1 & 1 \\ \cos\omega_1 & \cos\omega_2 \\ \cos 2\omega_1 & \cos 2\omega_2 \\ \cos 3\omega_1 & \cos 3\omega_2 \\ \cos 4\omega_1 & \cos 4\omega_2 \\ \cos 5\omega_1 & \cos 5\omega_2 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} 0.0691 \\ 0.0078 \\ -0.0019 \\ 0.0675 \\ 0.0188 \\ -0.0096 \end{bmatrix}$$

and the rank of the symmetric Toeplitz matrix is $M = 2$. For this matrix, we have

$$\|(T + UJ_{1/2})^{+}\| = 8.7306, \quad \|A\| = 0.1584, \quad \text{and} \quad \|A^{+}\| = 5.4734 \times 10^{6}.$$

Thus, $\kappa_0^{\mathcal{S}}$ is much smaller than $\kappa^{\mathcal{S}}$. Also, this example shows that by exploiting the structure, $\kappa_0^{\mathcal{S}}$ can be much smaller than κ_0.

Now, we remove the assumption $\mathbf{x} = J\mathbf{x}$. For a general n-vector solution $\mathbf{x}$, we partition

$$\mathbf{x} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix},$$

where $\mathbf{u}$ is an m-vector, $m = \lceil n/2 \rceil$. Denoting

$$\mathbf{y}_1 := \frac{1}{2}\big(\mathbf{u} + J_{1/2}^{\mathrm{T}}\mathbf{v}\big), \quad \mathbf{y} := \begin{bmatrix} \mathbf{y}_1 \\ J_{1/2}\mathbf{y}_1 \end{bmatrix},$$

$$\mathbf{z}_1 := \frac{1}{2}\big(\mathbf{u} - J_{1/2}^{\mathrm{T}}\mathbf{v}\big), \quad \mathbf{z} := \begin{bmatrix} \mathbf{z}_1 \\ -J_{1/2}\mathbf{z}_1 \end{bmatrix},$$

we have

$$J\mathbf{y} = \mathbf{y} \quad \text{and} \quad J\mathbf{z} = -\mathbf{z},$$

and

$$\mathbf{x} = \mathbf{y} + \mathbf{z},$$

noting that $J_{1/2}J_{1/2}^{\mathrm{T}} = I$. From the above argument, we have

$$\|A^+E\mathbf{x}\| = \|A^+E(\mathbf{y}+\mathbf{z})\| \le \big\|(T + UJ_{1/2})^{\#}\big\| \, \|\mathbf{y}\| + \big\|(T - UJ_{1/2})^{\#}\big\| \, \|\mathbf{z}\|.$$

7 Hankel.

For the Hankel structure, we have the following results.

THEOREM 7.1. *For the structure $\mathcal{S} = \{Hankel\}$,*

$$\kappa^{\text{Hankel}} \ge \sqrt{\frac{2\|A^+\| \, \|\mathbf{b}\|}{\|\mathbf{x}\|}}$$

and

$$\frac{\kappa^{\text{Hankel}}}{\kappa} \ge \frac{1}{\sqrt{2}\,\|A\| \, \|A^+\|}.$$

PROOF. Let $E = A/\|A\|$ in (2.8), we have

$$\varphi \ge \frac{\|A^+A\mathbf{x}\|}{\|A\|} = \frac{\|\mathbf{x}\|}{\|A\|},$$

which implies that

$$\varphi\,\|A\| + \|A^+\| \, \|\mathbf{b}\| \ge \|\mathbf{x}\| + \|A^+\| \, \|\mathbf{b}\| \ge 2\sqrt{\|A^+\| \, \|\mathbf{b}\| \, \|\mathbf{x}\|}.$$

Consequently, from (2.11) in Theorem 2.1,

$$\kappa^{\text{Hankel}} \ge \frac{1}{\sqrt{2}\,\|\mathbf{x}\|}(\varphi\,\|A\| + \|A^+\| \, \|\mathbf{b}\|) \ge \sqrt{\frac{2\|A^+\| \, \|\mathbf{b}\|}{\|\mathbf{x}\|}}.$$

Since $\|\mathbf{x}\| \leq \|A^+\|\,\|\mathbf{b}\|$, we have $\kappa^{\mathcal{S}} \geq \sqrt{2}$. Then it follows from (1.3) that

$$\frac{\kappa^{\text{Hankel}}}{\kappa} \geq \frac{1}{\sqrt{2}\,\|A\|\,\|A^+\|}.$$

□

Like the circulant case, this shows that we may expect a smaller condition number for Hankel least squares problems.

8 Generalized inverse.

Similar to the definition (2.1) of the condition number $\kappa^{\mathcal{S}}$ for structured least squares, we define

$$\kappa^{\mathcal{S}}(A) := \limsup \left\{ \frac{\|(A+\Delta A)^+ - A^+\|}{\epsilon\,\|A^+\|} \ :\ A, \Delta A \in \mathcal{S},\ \ \|\Delta A\| \leq \epsilon\,\|A\| \right\}$$

as the condition number for the generalized inversion of matrix A.

For ΔA satisfying (2.2), we have (2.3), which implies that

$$\|(A+\Delta A)^+ - A^+\| \approx \|A^+(\Delta A)A^+\| \leq \epsilon \|A^+\|^2\,\|A\|.$$

Thus, we have

$$\kappa^{\mathcal{S}}(A) \leq \|A^+\|\,\|A\|.$$

On the other hand, let $\Delta A = \epsilon A$, then ΔA and A have the same structure and $\|\Delta A\| = \epsilon\|A\|$. Since for this particular ΔA

$$\frac{\|A^+(\Delta A)A^+\|}{\epsilon\|A^+\|} = \frac{\|A^+AA^+\|}{\|A^+\|} = 1,$$

we have $\kappa^{\mathcal{S}}(A) \geq 1$. Thus, in general,

$$1 \leq \kappa^{\mathcal{S}}(A) \leq \|A^+\|\,\|A\|.$$

As expected, for $\mathcal{S} = \{\text{symmetric, skewsymmetric}\}$, the condition number is the same as the unstructured condition number

$$\kappa(A) = \|A^+\|\,\|A\|.$$

It turns out that this is also the same for circulant structure.

THEOREM 8.1. *For $\mathcal{S} = \{$symmetric, skewsymmetric, circulant$\}$,*

$$\kappa^{\mathcal{S}}(A) = \|A^+\|\,\|A\|.$$

PROOF. We show that $\kappa^{\mathcal{S}}(A) \geq \|A^+\|\,\|A\|$. Recall that the matrix E is defined by $\Delta A = \epsilon\|A\|E$ such that $E \in \mathcal{S}$ and $\|E\| \leq 1$. In the following, we show that there is an E such that $\|A^+EA^+\| \geq \|A^+\|^2$, which implies that

$$\frac{\|(A+\Delta A)^+ - A^+\|}{\epsilon\,\|A^+\|} \approx \frac{\|A^+(\Delta A)A^+\|}{\epsilon\,\|A^+\|} = \frac{\|A\|\,\|A^+EA^+\|}{\|A^+\|} \geq \|A^+\|\,\|A\|$$

for this particular E. It then follows that $\kappa^{\mathcal{S}}(A) \geq \|A^+\|\,\|A\|$.

Case 1 (symmetric). Let $A^+ = U\Sigma V^{\mathrm{T}}$ be the SVD of A^+ and $\mathbf{u}$ and $\mathbf{v}$ the singular vectors corresponding to the largest singular value of A^+, then $A^+\mathbf{v} = \|A^+\|\mathbf{u}$. From Lemma 3.1, there exists a symmetric E so that $E\mathbf{u} = \mathbf{v}$ and $\|E\| = 1$. Thus

$$\|A^+EA^+\mathbf{v}\| = \|A^+\|\,\|A^+E\mathbf{u}\| = \|A^+\|\,\|A^+\mathbf{v}\| = \|A^+\|^2$$

and we have $\|A^+EA^+\| \geq \|A^+\|^2$.

Case 2 (skewsymmetric). Let $A^+ = U\Sigma V^{\mathrm{T}}$ be the SVD of A^+. Suppose that $\mathbf{u}$ and $\mathbf{v}$ are the singular vectors such that $A^+\mathbf{v} = \|A^+\|\mathbf{u}$. Since A^+ is also skewsymmetric, we have $A^+ = -(A^+)^{\mathrm{T}} = -V\Sigma U^{\mathrm{T}}$. From this another form of the SVD of A^+, we have $A^+\mathbf{u} = -\|A^+\|\mathbf{v}$. Let $E = A^+/\|A^+\|$, then E is skewsymmetric, $\|E\| = 1$, and

$$\|A^+EA^+\mathbf{v}\| = \|A^+\|\,\|A^+E\mathbf{u}\| = \|A^+\|\,\|A^+\mathbf{v}\| = \|A^+\|^2.$$

Thus, $\|A^+EA^+\| \geq \|A^+\|^2$.

Case 3 (circulant). For a circulant matrix A, choosing $E = A^+/\|A^+\|$ and using the Fourier transformation diagonalization $A^+ = F^{-1}DF$, we get

$$\|A^+EA^+\| = \frac{\|(A^+)^3\|}{\|A^+\|} = \|A^+\|^2,$$

and E is circulant and $\|E\| = 1$. $\square$

Conclusion.

The condition number κ for the general least squares (LS) problem is a well-studied subject. This paper considers the case when the coefficient matrix is structured. In this case, it makes sense to require that the coefficient matrix and its perturbation matrix have the same structure. Based on this assumption, first we have defined and derived the condition number $\kappa^{\mathcal{S}}$ for the LS problems with structure $\mathcal{S}$. Then we have discussed special structures. For symmetric, skew symmetric, and persymmetric structures, we have shown that the condition numbers for structured and general LS problems are identical. For circulant, symmetric Toeplitz, and Hankel structures, we have presented a condition number $\kappa^{\mathcal{S}}$ for $\mathcal{S} = \{\text{circulant, symmetric Toeplitz, Hankel}\}$ and studied the condition number for each of the three structures. Our analysis has shown that $\kappa^{\mathcal{S}}$ can be smaller than κ when the structure $\mathcal{S}$ is circulant or symmetric Toeplitz or Hankel. Finally, the condition numbers for the related problem: generalized inversion of a structured matrix is also discussed. We have shown no improvement on the condition number for the generalized inverse when the structure $\mathcal{S} = \{\text{symmetric, skewsymmetric, circulant}\}$.

Acknowledgment.

The authors would like to thank Prof. S. M. Rump for providing the manuscripts of [8, 9] and the referee for valuable comments.

REFERENCES

1. A. Ben-Israel, *On error bounds for generalized inverses*, SIAM J. Numer. Anal., 3 (1966), pp. 585–592.
2. A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, 2nd edn., Springer, New York, 2003.
3. P. J. Davis, *Circulant Matrices*, John Wiley, New York, 1979.
4. G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd edn., Johns Hopkins University Press, Baltimore, MD, 1996.
5. D. J. Higham, *Condition numbers and their condition numbers*, Linear Algebra Appl., 214 (1995), pp. 193–213.
6. N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd edn., Society for Industrial and Applied Mathematics, Philadelphia, PA, 2002.
7. S. M. Rump, *Ill-conditionedness need not be the componentwise near to ill-posedness for least squares problems,* BIT, 39 (1999), pp. 143–151.
8. S. M. Rump, *Structured perturbations. Part I: Normwise distances,* SIAM J. Matrix Anal. Appl., 25 (2003), pp. 1–30.
9. S. M. Rump, *Structured perturbation. Part II: Componentwise distances*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 31–56.
10. G. Wang, Y. Wei, and S. Qiao, *Generalized Inverses: Theory and Computations*, Science Press, Beijing/New York, 2004.
11. Y. Wei and H. Wu, *Expressions for the perturbation of the weighted Moore–Penrose inverse,* Comput. Math. Appl., 39 (2000), pp. 13–18.
12. Y. Wei and D. Wang, *Condition numbers and perturbation of the weighted Moore–Penrose inverse and weighted least squares problem*, Appl. Math. Comput., 145 (2003), pp. 45–58.