Journal of Computational Mathematics, Vol.22, No.4, 2004, 605–610.

STRUCTURED BACKWARD ERRORS FOR STRUCTURED KKT SYSTEMS *1)

Xin-xiu Li
(*Department of Applied Mathematics and Physics, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*)

Xin-guo Liu
(*Department of Mathematics, Ocean University of Qingdao, Qingdao 266071, China*)

Abstract

In this paper we study structured backward errors for some structured KKT systems. Normwise structured backward errors for structured KKT systems are defined, and computable formulae of the structured backward errors are obtained. Simple numerical examples show that the structured backward errors may be much larger than the unstructured ones in some cases.

Mathematics subject classification: 65F25.
Key words: Structured — KKT system, Structured backward error, Normwise backward error.

1. Introduction

Consider the problem of solving the following structured linear systems

$$\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}, \tag{1}$$

for x and y, where $A \in \mathcal{R}^{m\times n}, x, b \in \mathcal{R}^m, y \in \mathcal{R}^n$;

$$\begin{pmatrix} I & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} B \\ 0 \end{pmatrix}, \tag{2}$$

for X and Y, where $A \in \mathcal{R}^{m\times n}, B, X \in \mathcal{R}^{m\times r}, Y \in \mathcal{R}^{n\times r}$;

$$\begin{pmatrix} 0 & 0 & B \\ 0 & I & A \\ B^T & A^T & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} d \\ b \\ 0 \end{pmatrix}, \tag{3}$$

and

$$\begin{pmatrix} 0 & 0 & C & B \\ 0 & I & 0 & I \\ C^T & 0 & 0 & 0 \\ B^T & I & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{4}$$

These systems called augmented systems, and they are structured Karush–Kuhn–Tucker (structured — KKT) systems. The structured – KKT systems of (1) – (4) arise in many applications, for example, for the linear least squares problem

$$\text{LS}: \quad \min_{y\in\mathcal{R}^n} \|b - Ay\|_2, \quad A \in \mathcal{R}^{m\times n}, \quad b \in \mathcal{R}^m,$$

let $r = b - Ay$, then the LS minimizer y satisfies the structured — KKT system (1) since this is simply a representation of the normal equations. The structured — KKT systems (2) – (4)

* Received June 27, 2002.
1) This work is supported by NNSF of China.

mainly arise in the linear least squares problem with multiple right-hand sides[19], the linear least squares problem with equality constraint[3,4], and the generalized linear least squares problem[14], respectively. And problems of least squares problems arise in many fields of study.

The structured linear systems (1) – (4) are structured — KKT systems, and they have stronger structure than the KKT systems, so the results of Sun[16] about the KKT systems are invalid to the structured linear systmes (1) – (4).Consequently, the structured optimal backward perturbation analysis of the structured — KKT systems (1) – (4) is worth researching.

Consider the linear system $Ax = b$. Let $\hat{x}$ be a computed solution to the system. In general, there are many perturbations ΔA, and Δb such that $\hat{x}$ is a solution to the perturbed systems $(A + \Delta A)x = b + \Delta b$. It may be asked how close is the nearest system for which $\hat{x}$ is the solution to the original system. There are various approaches to define backward errors (BEs) for measuring the distance between the perturbed systems and the original systems. Finding an explicit expression of a BE may be very useful for testing the stability of practical algorithms. For general linear system, Rigal and Gaches[15] have defined a normwise BE and obtained the explicit expression

$$\tau(\hat{x}) = \frac{\|b - A\hat{x}\|_2}{\sqrt{\|A\|_F^2\|\hat{x}\|_2^2 + \|b\|_2^2}}.$$

However, it is worth pointing out that if the coefficient matrix A has some special structure, and the perturbed matrice $A + \Delta A$ has the same form as A, in which we are interested. The problem of finding an expression of the corresponding BE should be concerned. Generally speaking, Rigal and Gaches' result is a strict lower bound of the structured BE.

In this paper we shall define the structured BEs of Eq. (1)-(4), derive computable formulae of them, and show that if the perturbed matrices have the same form as the coefficient matrix and the unstructured backward error $\tau(\hat{x})$ is small, it does not necessarily follow that $\hat{x}$ solves a nearby structured linear system.

2. Structured Backward Errors for Structured-KKT Systems

Firstly, we investigate the backward errors for structured-KKT systems (1).

Theorem 2.1. *Let $(\hat{x}^T, \hat{y}^T)^T$ with $\hat{y} \neq 0$ be a computed solution of Eq. (1). Define the normwise structured backward error $\eta^{(\theta)}(\hat{x}, \hat{y})$ of the Eq. (1) by*

$$\eta^{(\theta)}(\hat{x}, \hat{y}) = \min_{(\Delta A, \Delta b) \in \mathcal{E}} \|(\Delta A, \theta\Delta b)\|_F, \tag{5}$$

where θ is a positive parameter, and the set $\mathcal{E}$ is defined by

$$\mathcal{E} = \left\{ (\Delta A, \Delta b) : \begin{pmatrix} I & A + \Delta A \\ (A + \Delta A)^T & 0 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = \begin{pmatrix} b + \Delta b \\ 0 \end{pmatrix} \right\}. \tag{6}$$

Then

$$\eta^{(\theta)}(\hat{x}, \hat{y}) = \left[\hat{x}^\dagger (AA^T + \theta^2 (\hat{x} - b)(\hat{x} - b)^T)\hat{x} + \tau \|(I_m - \hat{x}\hat{x}^\dagger)(A\hat{y} - b)\|_2^2 \right]^{1/2}. \tag{7}$$

The corresponding perturbations of A and b are

$$\begin{aligned} \Delta A^* &= -\hat{x}\hat{x}^\dagger A + \tau (I_m - \hat{x}\hat{x}^\dagger)(b - A\hat{y})\hat{y}^T \quad , \\ \Delta b^* &= \hat{x}\hat{x}^\dagger (\hat{x} - b) - \frac{1}{1 + \theta^2\|\hat{y}\|_2^2}(I_m - \hat{x}\hat{x}^\dagger)(b - A\hat{y}) \ . \end{aligned} \tag{8}$$

That is $\eta^{(\theta)}(\hat{x}, \hat{y}) = \|(\Delta A^*, \theta\Delta b^*)\|_F$, where $\tau = \dfrac{\theta^2}{1 + \theta^2\|\hat{x}\|_2^2}$.

Proof. By Eq. (6), $(\Delta A, \Delta b) \in \mathcal{E}$ if and only if ΔA and Δb satisfy

$$\Delta b = \hat{x} - b + A\hat{y} + \Delta A\hat{y}, \tag{9}$$

$$\hat{x}^T \Delta A = -\hat{x}^T A. \tag{10}$$

From (10),

$$\Delta A = -\hat{x}\hat{x}^\dagger A + (I_m - \hat{x}\hat{x}^\dagger)Z, \quad Z \in \mathcal{R}^{m \times n}. \tag{11}$$