SIMULATION OPTIMIZATION SYSTEMS
Research Laboratory

A UNIFIED APPROACH TO FIRST-ORDER AND LARGE CHANGE SENSITIVITY COMPUTATIONS IN LINEAR SYSTEMS

J.W. Bandler and Q.J. Zhang

SOS-84-20-R

September 1984

McMASTER UNIVERSITY
Hamilton, Canada L8S 4L7
Department of Electrical and Computer Engineering

A UNIFIED APPROACH TO FIRST-ORDER AND LARGE CHANGE SENSITIVITY COMPUTATIONS IN LINEAR SYSTEMS

J.W. Bandler and Q.J. Zhang

SOS-84-20-R

September 1984

A UNIFIED APPROACH TO FIRST-ORDER AND LARGE CHANGE SENSITIVITY COMPUTATIONS IN LINEAR SYSTEMS

J.W. Bandler and Q.J. Zhang

Abstract

This paper presents a unified approach to sensitivity analysis in linear systems. A comprehensive set of first-order and large change sensitivity formulas associated with matrix and vector operations is provided. They can be directly applied to linear system simulation and optimization procedures. Different formulas are compared. Relevant expressions appropriate for computational purposes are listed. These schemes will be very powerful if, for any particular sensitivity analysis, the variables exist only in one or several local areas of the system. In an optimization procedure, the calculation of large change effects can be reduced from solving an n x n system to solving an r x r system, where r is the rank of the deviation matrix of the linear system.

This work was supported by the Natural Sciences and Engineering Research Council of Canada under Grant A7239.

The authors are with the Simulation Optimization System Research Laboratory and the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Canada L8S 4L7.

I. INTRODUCTION

Sensitivity analysis remains one of the important topics in both mathematical and engineering problems. In computer oriented simulation and optimization, it is often required to calculate the gradients of functions of interest w.r.t. variable parameters and to recalculate new function values when the values of certain variables are changed. Besides, sensitivity analysis itself can yield valuable information about the system under consideration. An elegant approach to the evaluation of these sensitivities should be such that calculations are performed efficiently using all possible information to reduce computational effort.

In this paper, we investigate first-order and large change sensitivity problems and formulate a comprehensive set of formulas associated with matrix and vector operations. They can be directly applied to linear systems. Different formulas are compared. Relevant expressions appropriate for computational purposes are provided. These formulas will be very powerful, especially when the variables exist only in some local areas of the linear system. In an optimization procedure, the calculation of large change effects can be reduced from solving an n x n system to solving an r x r system, where r is the rank of the deviation matrix of the linear system.

In Section II, we introduce the basic notation, definitions and relations which will be used throughout this paper. We first present the basic and relatively easier sensitivity formulas w.r.t. matrices, vectors or their components in Section III, followed by first-order sensitivity formulas w.r.t. arbitrary variables in Section IV. Section V is devoted to the derivation of formulas for evaluating large changes in $\mathbf{A}^{-1}$, which is further discussed in Section VI. Section VII shows the expansion of basic large change formulas to various expressions suitable for computational purposes due to different formulations of the original problem. Section VIII provides operational counts for major formulas to facilitate a deep investigation of the computational effort. Appropriate formulations yielding efficient calculation of large change effects is further discussed in Section IX. Numerical examples are given in Section X. A comprehensive set of expressions designed for computation are provided through tables which summarize and complete the various cases of our analysis. These tables can be used as a tool for computer-aided designers to yield efficient design procedures.

II. BASIC NOTATION, DEFINITIONS AND RELATIONS

Basic Notation

Let

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix} \quad (1)$$

be an m x n matrix containing the variables $\boldsymbol{\phi}$ which exist only in rows $i_1, i_2, \ldots, i_{nI}$ and columns $j_1, j_2, \ldots, j_{nJ}$. Let M, N, I and J be index sets defined by

$$M \triangleq \{1, 2, \ldots, m\}, \quad (2)$$

$$N \triangleq \{1, 2, \ldots, n\}, \quad (3)$$

$$I \triangleq \{i_1, i_2, \ldots, i_{nI}\} \subset M, \quad (4)$$

$$J \triangleq \{j_1, j_2, \ldots, j_{nJ}\} \subset N. \quad (5)$$

Let $\mathbf{U}_I$ and $\mathbf{U}_J$ be matrices whose columns are unit vectors $\mathbf{u}_{i_1}, \mathbf{u}_{i_2} \ldots, \mathbf{u}_{i_{nI}}$ and $\mathbf{u}_{j_1}, \mathbf{u}_{j_2}, \ldots, \mathbf{u}_{j_{nJ}}$, respectively, i.e.,

$$\mathbf{U}_I \triangleq [\mathbf{u}_{i_1} \ \mathbf{u}_{i_2} \ldots \mathbf{u}_{i_{nI}}], \quad (6)$$

$$\mathbf{U}_J \triangleq [\mathbf{u}_{j_1} \ \mathbf{u}_{j_2} \ldots \mathbf{u}_{j_{nJ}}], \quad (7)$$

where $\mathbf{u}_i$, $i \in I$, is an m-dimensional unit-vector with 1 in its ith row and zeros everywhere else. Similarly, $\mathbf{u}_j$, $j \in J$, is an n-dimensional unit vector with 1 in the jth row and zeros everywhere else. Vectors $\mathbf{u}_i$ and $\mathbf{u}_j$ are named unit vectors.

The intersection elements of rows $i_1, i_2, \ldots, i_{nI}$ and columns $j_1, j_2, \ldots, j_{nJ}$ in $\mathbf{A}$ constitute an n_I by n_J submatrix denoted by $\mathbf{A}_{IJ}$, namely,

$$\mathbf{A}_{IJ} = \begin{bmatrix} A_{i_1 j_1} & A_{i_1 j_2} & \cdots & A_{i_1 j_{nJ}} \\ A_{i_2 j_1} & A_{i_2 j_2} & \cdots & A_{i_2 j_{nJ}} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ A_{i_{nI} j_1} & A_{i_{nI} j_2} & \cdots & A_{i_{nI} j_{nJ}} \end{bmatrix} . \quad (8)$$

The matrices $\mathbf{U}_I$ and $\mathbf{U}_J$ have the property of selecting certain components of a matrix by premultiplication and postmultiplication. For example, we have the relation

$$\mathbf{A}_{IJ} = \mathbf{U}_I^T \mathbf{A} \, \mathbf{U}_J . \quad (9)$$

Basic Definition

Since $\mathbf{A}$ contains the variable $\boldsymbol{\phi}$, a large change in $\boldsymbol{\phi}$ will cause large changes in $\mathbf{A}$ and also in $\mathbf{A}^{-1}$ when m = n and $\mathbf{A}$ is nonsingular. Let $\Delta\mathbf{A}$ represent the large changes of $\mathbf{A}$ due to changes in $\boldsymbol{\phi}$. We define

$$\Delta(\mathbf{A}^{-1}) \triangleq (\mathbf{A} + \Delta\mathbf{A})^{-1} - \mathbf{A}^{-1} . \quad (10)$$

Basic Relation: The Product Form of $\Delta\mathbf{A}$

It should be noted that $\Delta\mathbf{A}$ can always be expressed in the form of

$$\Delta\mathbf{A} = \mathbf{V}\,\mathbf{D}\,\mathbf{W}^T, \quad (11)$$

where $\mathbf{V}$, $\mathbf{D}$ and $\mathbf{W}$ are matrices of order of $m \times r_1$, $r_1 \times r_2$ and $n \times r_2$, respectively. We have great freedom to choose $\mathbf{V}$, $\mathbf{D}$ and $\mathbf{W}$ for the realization of (11). For example, a simple and trivial realization is

$$\Delta\mathbf{A} = \mathbf{1}_{m \times m}\, \Delta\mathbf{A}\, \mathbf{1}_{n \times n}, \quad (12)$$

where $\mathbf{1}_{m \times m}$ and $\mathbf{1}_{n \times n}$ are identity matrices of orders m and n, respectively. Another realization, which is very useful, is

$$\Delta\mathbf{A} = \mathbf{U}_I\, \Delta\mathbf{A}_{IJ}\, \mathbf{U}_J^T . \quad (13)$$

This equation is evident since, according to our previous assumption, the variables $\boldsymbol{\phi}$ exist in $\mathbf{A}$ only in elements A_{ij}, $i \in I, j \in J$.

We can fully exploit the freedom of selecting **V**, **D** and **W** so that various efficient computational schemes are constructed. This will be further discussed in Section IX.

Table I summarizes all basic notation used in the paper. Temporary notation is explained where it is used.

Representation of Solutions of Linear Equations

P or **p** and **Q** or **q** are used to represent the solutions of the linear systems with coefficient matrix **A** and $\mathbf{A}^T$, respectively, i.e.,

$$\mathbf{A\,P} = \mathbf{RHS}\,, \tag{14}$$

$$\mathbf{A\,p} = \mathbf{RHS}\,, \tag{15}$$

$$\mathbf{A}^T \mathbf{Q} = \mathbf{RHS}\,, \tag{16}$$

$$\mathbf{A}^T \mathbf{q} = \mathbf{RHS}\,, \tag{17}$$

where **RHS** is a generic R.H.S. vector or matrix of consistent dimensions with **P**, **p**, **Q** and **q**.

Subscripts similar to the notation of R.H.S. are added to **P**, **p**, **Q** and **q** to identify the solutions for different R.H.S. as shown in Table II. As an example, $\mathbf{p}_b$ and $\mathbf{Q}_W$ are solutions of

$$\mathbf{A\,p}_b = \mathbf{b} \tag{18}$$

and

$$\mathbf{A}^T \mathbf{Q}_W = \mathbf{W}\,, \tag{19}$$

respectively.

The linear equations with coefficient matrix **A** (e.g., equations (14), (15), (18)) and $\mathbf{A}^T$ (e.g., (16), (17), (19)) will be called the original and the adjoint systems of **A**, respectively.

III. SENSITIVITY FORMULAS W.R.T. MATRICES, VECTORS OR THEIR COMPONENTS

In this section, we discuss first-order sensitivities w.r.t. variables that are the components of a matrix or a vector [1]. A complete set of these formulas and associated computational expressions are provided in Tables III, IV and V. Table IV gives relevant formulas for the case when **A** is symmetrical.

Sensitivities of A and $\mathbf{A}^{-1}$

To begin with, we can easily find that the sensitivity of matrix **A** w.r.t. its (i,j) component is a matrix with 1 at the (i,j) position and zeros everywhere else. Using unit vectors $\mathbf{u}_i$ and $\mathbf{u}_j$, this sensitivity can be expressed as

$$\frac{\partial \mathbf{A}}{\partial A_{ij}} = \mathbf{u}_i \, \mathbf{u}_j^T . \tag{20}$$

By differentiating $\mathbf{A}\,\mathbf{A}^{-1} = \mathbf{1}$ and manipulating relevant terms [1], we have

$$\frac{\partial \mathbf{A}^{-1}}{\partial \phi} = - \mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \phi} \mathbf{A}^{-1} . \tag{21}$$

Letting $\phi = A_{ij}$ and substituting (20) into (21), we have

$$\frac{\partial \mathbf{A}^{-1}}{\partial A_{ij}} = - \mathbf{A}^{-1} \mathbf{u}_i \, \mathbf{u}_j^T \mathbf{A}^{-1} . \tag{22}$$

The expression appropriate for the computation of (22) is

$$\frac{\partial \mathbf{A}^{-1}}{\partial A_{ij}} = - \mathbf{p}_{ui} \, \mathbf{q}_{uj}^T , \tag{23}$$

where $\mathbf{p}_{ui}$ and $\mathbf{q}_{uj}$ are the solutions of the original and adjoint linear systems of **A** with R.H.S. as $\mathbf{u}_i$ and $\mathbf{u}_j$, respectively, as defined in Table II. The major computational effort involved in (23) is one LU factorization and two forward and backward substitutions.

Sensitivities of ($\mathbf{B}^T \mathbf{A}\,\mathbf{C}$)

The sensitivities of the matrix product $\mathbf{B}^T\mathbf{A}\,\mathbf{C}$ w.r.t. A_{ij} can be derived, using equation (20) as

$$\frac{\partial (\mathbf{B}^T \mathbf{A}\,\mathbf{C})}{\partial A_{ij}} = \mathbf{B}^T \frac{\partial \mathbf{A}}{\partial A_{ij}} \mathbf{C}$$

$$= \mathbf{B}^T \mathbf{u}_i \, \mathbf{u}_j^T \mathbf{C} . \tag{24}$$

The sensitivities of the (ℓ, k) component of $(\mathbf{B}^T \mathbf{A}\, \mathbf{C})$ (denoted $[\mathbf{B}^T \mathbf{A}\, \mathbf{C}]_{\ell k}$) w.r.t. matrix $\mathbf{A}$ is a matrix the same order as $\mathbf{A}$. The (i, j) component of the sensitivity matrix is

$$\frac{\partial [\mathbf{B}^T \mathbf{A}\,\mathbf{C}]_{\ell k}}{\partial A_{ij}} = \frac{\partial (\mathbf{u}_\ell^T \mathbf{B}^T \mathbf{A}\,\mathbf{C}\,\mathbf{u}_k)}{\partial A_{ij}}$$

$$= \mathbf{u}_\ell^T \mathbf{B}^T \frac{\partial \mathbf{A}}{\partial A_{ij}} \mathbf{C}\,\mathbf{u}_k$$

$$= (\mathbf{u}_\ell^T \mathbf{B}^T \mathbf{u}_i)(\mathbf{u}_j^T\, \mathbf{C}\,\mathbf{u}_k)$$

$$= (\mathbf{u}_i^T \mathbf{B}\,\mathbf{u}_\ell)(\mathbf{u}_k^T\, \mathbf{C}^T \mathbf{u}_j)$$

$$= \mathbf{u}_i^T (\mathbf{B}\,\mathbf{u}_\ell \mathbf{u}_k^T\, \mathbf{C}^T)\,\mathbf{u}_j , \tag{25}$$

where $\mathbf{u}_\ell$ and $\mathbf{u}_k$ are unit vectors with appropriate dimensions for the postmultiplication to $\mathbf{B}$ and $\mathbf{C}$, respectively. Notice that equation (20) was used to substitute $\partial\mathbf{A}/\partial A_{ij}$. The R.H.S. of the last equality indicates that the result of the calculation of (25) is the (i,j) component of ($\mathbf{B}$ $\mathbf{u}_\ell\, \mathbf{u}_k{}^T\, \mathbf{C}^T$). Therefore, we find

$$\frac{\partial [\mathbf{B}^T \mathbf{A}\,\mathbf{C}]_{\ell k}}{\partial \mathbf{A}} = \mathbf{B}\,\mathbf{u}_\ell\; \mathbf{u}_k^T \mathbf{C}^T . \tag{26}$$

Comparing (26) with (24), we find that if $\mathbf{B}$ and $\mathbf{C}$ are both symmetrical matrices, then

$$\frac{\partial [\mathbf{B}^T \mathbf{A}\,\mathbf{C}]}{\partial A_{ij}} = \frac{\partial [\mathbf{B}^T \mathbf{A}\,\mathbf{C}]_{ij}}{\partial \mathbf{A}} . \tag{27}$$

<u>Sensitivities of $(\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})$</u>

Applying (22), we have

$$\frac{\partial (\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})}{\partial A_{ij}} = \mathbf{B}^T \frac{\partial\, \mathbf{A}^{-1}}{\partial A_{ij}} \mathbf{C}$$

$$= -\, \mathbf{B}^T \mathbf{A}^{-1} \mathbf{u}_i\, \mathbf{u}_j^T \mathbf{A}^{-1}\, \mathbf{C} . \tag{28}$$

Consider the (ℓ, k) component of the matrix product $\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}$ in terms of A_{ij} as

$$\frac{\partial [\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}]_{\ell k}}{\partial A_{ij}} = \frac{\partial (\mathbf{u}_\ell^T \mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}\, \mathbf{u}_k)}{\partial A_{ij}}$$

$$= \mathbf{u}_\ell^T \mathbf{B}^T \frac{\partial \mathbf{A}^{-1}}{\partial A_{ij}} \mathbf{C}\, \mathbf{u}_k . \tag{29}$$

Substituting (22) into (29), we find

$$\frac{\partial [\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}]_{\ell k}}{\partial A_{ij}} = -(\mathbf{u}_\ell^T \mathbf{B}^T \mathbf{A}^{-1} \mathbf{u}_i)(\mathbf{u}_j^T \mathbf{A}^{-1} \mathbf{C}\, \mathbf{u}_k)$$

$$= -\mathbf{u}_i^T [(\mathbf{A}^{-1})^T \mathbf{B}\, \mathbf{u}_\ell \mathbf{u}_k^T\, \mathbf{C}^T (\mathbf{A}^{-1})^T]\, \mathbf{u}_j . \tag{30}$$

Therefore, we have the sensitivities of the (ℓ, k) component of $\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}$ w.r.t. matrix $\mathbf{A}$ as

$$\frac{\partial [\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}]_{\ell k}}{\partial \mathbf{A}} = -(\mathbf{A}^{-1})^T \mathbf{B}\, \mathbf{u}_\ell \mathbf{u}_k^T\, \mathbf{C}^T (\mathbf{A}^{-1})^T . \tag{31}$$

Compare (31) with (28). We find that if matrix products $(\mathbf{B}^T \mathbf{A}^{-1})$ and $(\mathbf{A}^{-1} \mathbf{C})$ are both symmetrical, then

$$\frac{\partial (\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})}{\partial A_{ij}} = \frac{\partial [\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}]_{ij}}{\partial \mathbf{A}} . \tag{32}$$

Remark

Using these basic equations, various sensitivity expressions can be derived. Tables III and IV give a list of these expressions. Table V provides sensitivity expressions appropriate for computations when $\mathbf{A}^{-1}$ is involved. One LU-factorization and two forward and backward substitutions (FBS) are the major computations required for each of the sensitivity expressions in Table V.

IV. FIRST-ORDER SENSITIVITIES W.R.T. AN ARBITRARY VARIABLE

In the previous section, we considered the sensitivities w.r.t. matrix or vector components of a linear system. These components, however, may be functions of other variables, e.g., variable ϕ. Further, several components in the matrix or the vector may contain the same variable ϕ simultaneously. In this case, the chain rule may be used to obtain sensitivities of the functions of interest w.r.t. variable ϕ. For example, suppose $\mathbf{A} = \mathbf{A}(\phi)$, then

$$\frac{\partial \mathbf{A}}{\partial \phi} = \sum_{i,j} \left(\frac{\partial \mathbf{A}}{\partial A_{ij}} \frac{\partial A_{ij}}{\partial \phi} \right), \tag{33}$$

where the summation is performed over those indices of i and j corresponding to which the component A_{ij} contains the variable ϕ.

A more efficient approach can be developed when we consider simultaneously all the components affected by ϕ. As introduced in Section II, suppose ϕ exists in $\mathbf{A}$ in rows i, $i \in I$, and columns j, $j \in J$. Then we have

$$\frac{\partial \mathbf{A}}{\partial \phi} = \mathbf{U}_I \frac{\partial \mathbf{A}_{IJ}}{\partial \phi} \mathbf{U}_J^T. \tag{34}$$

The sensitivity of the inverse of $\mathbf{A}$ w.r.t. ϕ can be found as

$$\frac{\partial \mathbf{A}^{-1}}{\partial \phi} = -\mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial \phi} \mathbf{A}^{-1}$$

$$= -\mathbf{A}^{-1} \mathbf{U}_I \frac{\partial \mathbf{A}_{IJ}}{\partial \phi} \mathbf{U}_J^T \mathbf{A}^{-1}. \tag{35}$$

The corresponding expression suitable for computation, when $n_I + n_J < n$, is

$$\frac{\partial \mathbf{A}^{-1}}{\partial \phi} = -\mathbf{P}_{UI} \frac{\partial \mathbf{A}_{IJ}}{\partial \phi} \mathbf{Q}_{UJ}^T, \tag{36}$$

where $\mathbf{P}_{UI}$ and $\mathbf{Q}_{UJ}$ are solutions of the original and adjoint systems of $\mathbf{A}$ with R.H.S. $\mathbf{U}_I$ and $\mathbf{U}_J$, respectively.

When $n_I + n_J \geq n$, it is advantageous to perform a matrix inversion directly for $\mathbf{A}^{-1}$.

Using these basic formulas, various sensitivity expressions can be derived. Table VI shows sensitivity formulas suitable for calculations for different situations.

V. LARGE CHANGES OF A^{-1} DUE TO ARBITRARY VARIABLES IN A

Introduction

Consider the matrix $\mathbf{A}$ with arbitrary variables $\boldsymbol{\phi}$ which exist in components of $\mathbf{A}$ at rows i, $i \in I$, and columns j, $j \in J$. Suppose $\boldsymbol{\phi}$ is changed from $\boldsymbol{\phi}$ to $\boldsymbol{\phi} + \Delta\boldsymbol{\phi}$, then the changes of $\mathbf{A}$ can be easily found as

$$\Delta \mathbf{A} = \mathbf{U}_I \, \Delta \mathbf{A}_{IJ} \, \mathbf{U}_J^T . \tag{37}$$

The inverse of $\mathbf{A}$ will also be affected by the changes of $\boldsymbol{\phi}$. After the change, the inverse of $\mathbf{A}$ becomes $(\mathbf{A} + \Delta\mathbf{A})^{-1}$ which, according to our definition in (10), can be evaluated by

$$(\mathbf{A} + \Delta \mathbf{A})^{-1} = \mathbf{A}^{-1} + \Delta(\mathbf{A}^{-1}) , \tag{38}$$

where $\mathbf{A}^{-1}$ and $\Delta(\mathbf{A}^{-1})$ may be considered as the nominal value and incremental changes of matrix $\mathbf{A}^{-1}$.

Derivation of Major Large Change Formulas

Now, we find $\Delta(\mathbf{A}^{-1})$ as follows. Since

$$(\mathbf{A} + \Delta \mathbf{A})^{-1} (\mathbf{A} + \Delta \mathbf{A}) = \mathbf{1}, \tag{39}$$

substituting (38) into (39), we have

$$\mathbf{A}^{-1} \mathbf{A} + \mathbf{A}^{-1} \Delta \mathbf{A} + \Delta(\mathbf{A}^{-1}) (\mathbf{A} + \Delta \mathbf{A}) = \mathbf{1} . \tag{40}$$

The term $\mathbf{A}^{-1}\mathbf{A}$ is cancelled by the identity matrix at the right hand side. Thus, we can have

$$\begin{aligned} \Delta(\mathbf{A}^{-1}) &= - \mathbf{A}^{-1} \Delta \mathbf{A} (\mathbf{A} + \Delta \mathbf{A})^{-1} \\ &= - \mathbf{A}^{-1} \Delta \mathbf{A} (\mathbf{A}^{-1} + \Delta(\mathbf{A}^{-1})) . \end{aligned} \tag{41}$$

As discussed in Section II, $\Delta\mathbf{A}$ can be represented by the matrix product form of $\mathbf{V}\,\mathbf{D}\,\mathbf{W}^T$. Therefore, equation (41) becomes:

$$\begin{aligned} \Delta(\mathbf{A}^{-1}) &= - \mathbf{A}^{-1} \mathbf{V} \mathbf{D} \mathbf{W}^T (\mathbf{A}^{-1} + \Delta(\mathbf{A}^{-1})) \\ &= - \mathbf{A}^{-1} \mathbf{V} \mathbf{D} \mathbf{W}^T \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{V} \mathbf{D} \mathbf{W}^T \Delta(\mathbf{A}^{-1}) . \end{aligned} \tag{42}$$

Premultiplying (42) by $\mathbf{W}^T$ and denoting $\mathbf{Z}$ as

$$\mathbf{Z} = \mathbf{W}^T \Delta(\mathbf{A}^{-1}) , \tag{43}$$

we have

$$\mathbf{Z} = -\mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D} \mathbf{W}^T \mathbf{A}^{-1} - \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D} \mathbf{Z} \tag{44}$$

and consequently, **Z** can be solved from (44) as

$$\mathbf{Z} = -(1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D} \mathbf{W}^T \mathbf{A}^{-1}. \tag{45}$$

Since $\mathbf{Z} = \mathbf{W}^T \Delta(\mathbf{A}^{-1})$, the product form of $\mathbf{W}^T \Delta(\mathbf{A}^{-1})$ on the R.H.S. of (42) can be replaced by the R.H.S. of (45), i.e.,

$$\begin{aligned}
\Delta(\mathbf{A}^{-1}) = & -\mathbf{A}^{-1} \mathbf{V} \mathbf{D} \mathbf{W}^T \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{V} \mathbf{D} (1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} \\
& \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D} \mathbf{W}^T \mathbf{A}^{-1} \\
= & -\mathbf{A}^{-1} \mathbf{V} \mathbf{D} [1 - (1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D}] \mathbf{W}^T \mathbf{A}^{-1} \\
= & -\mathbf{A}^{-1} \mathbf{V} \mathbf{D} [(1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} (1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D}) \\
& -(1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D}] \mathbf{W}^T \mathbf{A}^{-1} \\
= & -\mathbf{A}^{-1} \mathbf{V} \mathbf{D} [(1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} + (1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D} \\
& -(1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D}] \mathbf{W}^T \mathbf{A}^{-1} \\
= & -\mathbf{A}^{-1} \mathbf{V} \mathbf{D} (1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} \mathbf{W}^T \mathbf{A}^{-1}.
\end{aligned} \tag{46}$$

Therefore, we have the following formula to calculate the large change effects of $\mathbf{A}^{-1}$ as

$$\Delta(\mathbf{A}^{-1}) = -\mathbf{A}^{-1} \mathbf{V} \mathbf{D} (1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})^{-1} \mathbf{W}^T \mathbf{A}^{-1}. \tag{47}$$

Similarly, we can obtain

$$\Delta(\mathbf{A}^{-1}) = -\mathbf{A}^{-1} \mathbf{V} (1 + \mathbf{D} \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V})^{-1} \mathbf{D} \mathbf{W}^T \mathbf{A}^{-1}. \tag{48}$$

Notice that **D** is a matrix of dimensions $r_1 \times r_2$. The inversions of $(1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})$ in (47) and $(1 + \mathbf{D} \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V})$ in (48) are the orders of r_2 and r_1, respectively.

Discussion

Equations (47) and (48) indicate that if we already know $\mathbf{A}^{-1}$, the large change effects of $\mathbf{A}^{-1}$ due to changes in elements of **A** may be evaluated by inverting an $r_1 \times r_1$ or $r_2 \times r_2$ matrix and performing appropriate matrix multiplications. To obtain $(\mathbf{A} + \Delta\mathbf{A})^{-1}$, we simply add $\Delta(\mathbf{A}^{-1})$ to $\mathbf{A}^{-1}$.

We can always choose **V**, **D**, and **W** such that (11) holds and $r_1 \leq n$ and $r_2 \leq n$. Especially, when rank $\Delta\mathbf{A} < n$, we can have $r_1 < n$ and/or $r_2 < n$. Thus, instead of directly inverting an n x n system, we can obtain the $(\mathbf{A} + \Delta\mathbf{A})^{-1}$ matrix by solving a smaller system of $r_1 \times r_1$ or $r_2 \times r_2$. This is one of the major significances of (47) and (48). Further discussion on the formulation of **V**, **D** and **W** as well as computational aspects of large change formulas associated with $\mathbf{A}^{-1}$ will be presented in Section IX.

VI. IMPORTANT SPECIAL CASES OF LARGE CHANGE FORMULAS FOR $\mathbf{A}^{-1}$

Case 1: Express $\Delta\mathbf{A}$ Directly by Unit Vectors Selecting Appropriate Components of $\Delta\mathbf{A}$

As discussed in (37), $\Delta\mathbf{A}$ can be expressed by a submatrix $\Delta\mathbf{A}_{IJ}$, which contains all the components of **A** affected by large changes of variables $\boldsymbol{\phi}$, and unit vectors which select $\Delta\mathbf{A}_{IJ}$. Comparing (37) with (11), we have

$$\Delta\mathbf{A} = \mathbf{V}\,\mathbf{D}\,\mathbf{W}^T = \mathbf{U}_I\,\Delta\mathbf{A}_{IJ}\,\mathbf{U}_J^{\,T}. \tag{49}$$

In this way, matrices **V**, **D** and **W** can be represented by $\mathbf{U}_I$, $\Delta\mathbf{A}_{IJ}$ and $\mathbf{U}_J$, respectively. Thus, (47) and (48) become

$$\Delta(\mathbf{A}^{-1}) = -\,\mathbf{A}^{-1}\,\mathbf{U}_I\,\Delta\mathbf{A}_{IJ}\,(\mathbf{1} + \mathbf{U}_J^{\,T}\,\mathbf{A}^{-1}\,\mathbf{U}_I\,\Delta\mathbf{A}_{IJ})^{-1}\,\mathbf{U}_J^{\,T}\,\mathbf{A}^{-1} \tag{50}$$

and

$$\Delta(\mathbf{A}^{-1}) = -\,\mathbf{A}^{-1}\,\mathbf{U}_I\,(\mathbf{1} + \Delta\mathbf{A}_{IJ}\,\mathbf{U}_J^{\,T}\,\mathbf{A}^{-1}\,\mathbf{U}_I)^{-1}\,\Delta\mathbf{A}_{IJ}\,\mathbf{U}_J^{\,T}\,\mathbf{A}^{-1}, \tag{51}$$

respectively. Using these formulas to obtain $\Delta\mathbf{A}^{-1}$, we need to perform $n_I + n_J$ forward and backward substitutions for $\mathbf{A}^{-1}\,\mathbf{U}_I$ and $\mathbf{U}_J^{\,T}\,\mathbf{A}^{-1}$ and to solve an $n_I \times n_I$ or $n_J \times n_J$ smaller system.

This special case suggests that the simplest way to choose **V**, **D** and **W** for formulas (47) and (48) is to use $\mathbf{U}_I$, $\Delta\mathbf{A}_{IJ}$ and $\mathbf{U}_J$.

Case 2: Householder Formulas [2, 3]

Let $\Delta\mathbf{A}$ be expressed by

$$\Delta\mathbf{A} = \mathbf{V}\,\mathbf{D}\,\mathbf{W}^T \tag{52}$$

such that **D** is square and nonsingular. Then,

$$\mathbf{D}(1 + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1} = \mathbf{D}[(\mathbf{D}^{-1} + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V})\,\mathbf{D}]^{-1}$$

$$= (\mathbf{D}^{-1} + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V})^{-1} \tag{53}$$

and

$$\mathbf{D}(1 + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1} = \mathbf{D}(1 + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1}\mathbf{D}^{-1}\mathbf{D}$$

$$= \mathbf{D}(\mathbf{D} + \mathbf{D}\,\mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1}\mathbf{D}. \tag{54}$$

Therefore, if **D** is an invertible square matrix, equation (47) becomes,

$$\Delta(\mathbf{A}^{-1}) = -\,\mathbf{A}^{-1}\mathbf{V}(\mathbf{D}^{-1} + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V})^{-1}\mathbf{W}^T\mathbf{A}^{-1} \tag{55}$$

and

$$\Delta(\mathbf{A}^{-1}) = -\,\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D}(\mathbf{D} + \mathbf{D}\,\mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1}\mathbf{D}\,\mathbf{W}^T\mathbf{A}^{-1} \tag{56}$$

or, alternatively, from our definition of $\Delta(\mathbf{A}^{-1})$ in (10),

$$(\mathbf{A} + \Delta\mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{V}(\mathbf{D}^{-1} + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V})^{-1}\mathbf{W}^T\mathbf{A}^{-1} \tag{57}$$

and

$$(\mathbf{A} + \Delta\mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{V}\,\mathbf{D}(\mathbf{D} + \mathbf{D}\,\mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1}\mathbf{D}\,\mathbf{W}^T\mathbf{A}^{-1}. \tag{58}$$

Formulas (57) and (58) are the well-known Householder formulas [2, 3].

It should be noted that, compared with the general formulas (47) or (48), the Householder formulas (57) or (55) require one additional matrix inversion for $\mathbf{D}^{-1}$ whereas (58) or (56) require more matrix multiplcations. Also, using the Householder formulas, $\Delta\mathbf{A}$ must be expressed in the form $\mathbf{V}\,\mathbf{D}\,\mathbf{W}^T$ such that **D** is a nonsingular square matrix.

Case 3: Rank 1 Change [4-6]

Suppose $\Delta\mathbf{A}$ is a rank 1 matrix expressed by

$$\Delta\mathbf{A} = \mathbf{v}\,\Delta\phi\,\mathbf{w}^T, \tag{59}$$

where **v** and **w** are n-vectors. In this case, matrices **V**, **D** and **W** in (11) become **v**, $\Delta\phi$ and **w**, respectively. Equation (47) becomes

$$\Delta(\mathbf{A}^{-1}) = -\frac{\mathbf{A}^{-1}\mathbf{v}\,\mathbf{w}^T\mathbf{A}^{-1}\Delta\phi}{1 + \mathbf{w}^T\mathbf{A}^{-1}\mathbf{v}\,\Delta\phi} \tag{60}$$

or, alternatively, using definition (10)

$$(\mathbf{A} + \Delta\mathbf{A})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{v}\,\mathbf{w}^T\mathbf{A}^{-1}\,\Delta\phi}{1 + \mathbf{w}^T\mathbf{A}^{-1}\mathbf{v}\,\Delta\phi}\,. \tag{61}$$

Equation (60) or (61) are the results of one of the pioneer works in this area by Sherman and Morrison [4, 5].

Although Sherman and Morrison's formula is derived for rank 1 changes in $\mathbf{A}$, it can be used iteratively to calculate general large change effects of $\mathbf{A}^{-1}$ for ranks of $\Delta\mathbf{A}$ higher than 1. For example, consider $\Delta\mathbf{A} = \mathbf{U}_I\,\Delta\mathbf{A}_{IJ}\,\mathbf{U}_J^T$ where large changes occur in columns j_1, $j_2, \ldots, j_{nJ}$. At the first iteration, only the j_1th column is changed, all other columns keep their old values. At the second iteration, changes occur only in the j_2th column while the j_1th column keeps its new value and the rest of the columns keep their old values. This procedure is continued for columns $j_3, j_4, \ldots, j_{nJ}$. The final $(\mathbf{A} + \Delta\mathbf{A})^{-1}$ is the inverse of $\mathbf{A}^{-1}$ after change. The procedure can be generally expressed by

$$(\mathbf{A} + \Delta\mathbf{A})^{-1} = \mathbf{R}_{r_2}\,, \tag{62}$$

where $\mathbf{R}_{r_2}$ is the result of an iterative procedure initialized by

$$\mathbf{R}_0 = \mathbf{A}^{-1} \tag{63}$$

and calculated, e.g., if $r_1 > r_2$, by

$$\left\{\begin{aligned} &\mathbf{s}_t = \mathbf{R}_{t-1}\,\mathbf{V}\,\mathbf{D}\,\mathbf{u}_t \\ &\mathbf{w}_t = \mathbf{W}\,\mathbf{u}_t \\ &\mathbf{R}_t = \mathbf{R}_{t-1} - \frac{\mathbf{s}_t\,\mathbf{w}_t^T\,\mathbf{R}_{t-1}}{1 + \mathbf{w}_t^T\,\mathbf{s}_t} \\ &\quad t = 1, 2, \ldots, r_2\,, \end{aligned}\right. \tag{64}$$

where $\mathbf{u}_t$ is a unit r_2-vector and where $\mathbf{V}\;\mathbf{D}\;\mathbf{W}^T$ is used instead of $\mathbf{U}_I\,\Delta\mathbf{A}_{IJ}\,\mathbf{U}_J^T$ as a general form of $\Delta\mathbf{A}$. A similar procedure exists for the case when $r_1 < r_2$. However, those approaches may cause more computational effort than previous formulas. A detailed discussion is provided in Section VIII.

The above different cases are summarized in Table VII. Other special cases, e.g., those from [7], are also possible from (47). These formulas require the solution of a new system which has the same size as the original n x n system.

VII. LARGE CHANGE FORMULAS AND THEIR COMPUTATION

Formulas When Matrix Inversion is Not Involved

Suppose variables $\boldsymbol{\phi}$ exist only in matrix $\mathbf{A}$ in rows i, $i \in I$, and columns j, $j \in J$. Evidently, the following relations exist:

$$\Delta(\mathbf{b}^T \mathbf{A}) = \mathbf{b}^T \mathbf{U}_I \Delta \mathbf{A}_{IJ} \mathbf{U}_J{}^T$$

$$= \mathbf{b}^T \mathbf{V} \mathbf{D} \mathbf{W}^T, \tag{65}$$

$$\Delta(\mathbf{A}\,\mathbf{c}) = \mathbf{U}_I \Delta \mathbf{A}_{IJ} \mathbf{U}_J{}^T \mathbf{c}$$

$$= \mathbf{V} \mathbf{D} \mathbf{W}^T \mathbf{c}, \tag{66}$$

$$\Delta(\mathbf{b}^T \mathbf{A}\,\mathbf{c}) = \mathbf{b}^T \mathbf{U}_I \Delta \mathbf{A}_{IJ} \mathbf{U}_J{}^T \mathbf{c}$$

$$= \mathbf{b}^T \mathbf{V} \mathbf{D} \mathbf{W}^T \mathbf{c} \tag{67}$$

and

$$\Delta(\mathbf{B}^T \mathbf{A}\,\mathbf{C}) = \mathbf{B}^T \mathbf{U}_I \Delta \mathbf{A}_{IJ} \mathbf{U}_J{}^T \mathbf{C}$$

$$= \mathbf{B}^T \mathbf{V} \mathbf{D} \mathbf{W}^T \mathbf{C}. \tag{68}$$

Formulas When $\mathbf{A}^{-1}$ is Involved

Using any of (47), (48), (50), (51), (55) and (56), we can obtain corresponding large change formulas for the inverse of $\mathbf{A}^{-1}$ associated with vectors $\mathbf{b}$ and $\mathbf{c}$ or matrices $\mathbf{B}$ and $\mathbf{C}$. For example, using (47), we have

$$\Delta(\mathbf{A}^{-1}\mathbf{c}) = -\,\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D}\,(1 + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1}\,\mathbf{W}^T\mathbf{A}^{-1}\mathbf{c} \tag{69}$$

and

$$\Delta(\mathbf{B}^T\mathbf{A}^{-1}\mathbf{C}) = -\,\mathbf{B}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D}\,(1 + \mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1}\,\mathbf{W}^T\mathbf{A}^{-1}\mathbf{C}. \tag{70}$$

Discussion on the Computational Aspects of (69) and (70)

From the computational point of view, (69) can be expressed as

$$\Delta(\mathbf{A}^{-1}\,\mathbf{c}) = -\,\mathbf{P}_V\,\mathbf{D}\,\mathbf{s}\,, \tag{71}$$

where **s** is the solution of the $r_2 \times r_2$ linear system:

$$(\mathbf{1} + \mathbf{W}^T\,\mathbf{P}_V\,\mathbf{D})\,\mathbf{s} = \mathbf{W}^T\,\mathbf{p}_c\,. \tag{72}$$

As for (70), we first notice that the term $\mathbf{B}^T\,\mathbf{A}^{-1}\,\mathbf{V}$ can be computed either as $\mathbf{B}^T\,\mathbf{P}_V$ or ${\mathbf{Q}_B}^T\,\mathbf{V}$, where $\mathbf{P}_V$ and $\mathbf{Q}_B$ are solutions of the original and adjoint system of **A** with right-hand sides as **V** and **B**, respectively, and have been defined in Table II. The difference of operational counts [2] between $\mathbf{B}^T\,\mathbf{P}_V$ and ${\mathbf{Q}_B}^T\,\mathbf{V}$ is $n^2\,(r_1 - m')$ multiplication and divisions, where r_1 is the number of columns in **V**, m' is the number of columns in **B** and n is the order of the square matrix **A**, as defined in Table I. Therefore, comparing r_1 and m', $(\mathbf{B}^T\,\mathbf{A}^{-1}\,\mathbf{V})$ can be calculated by

$$\mathbf{B}^T\,\mathbf{A}^{-1}\,\mathbf{V} = \begin{cases} \mathbf{B}^T\,\mathbf{P}_V\,, & \text{if } r_1 \le m' \quad (73a) \\ \mathbf{Q}_B^T\,\mathbf{V}\,, & \text{if } r_1 > m'\,. \quad (73b) \end{cases}$$

Similarly, we find

$$\mathbf{W}^T\,\mathbf{A}^{-1}\,\mathbf{C} = \begin{cases} \mathbf{W}^T\,\mathbf{P}_C\,, & \text{if } r_2 > n' \quad (74a) \\ \mathbf{Q}_W^T\,\mathbf{C}\,, & \text{if } r_2 \le n'\,, \quad (74b) \end{cases}$$

where r_2 and n' are numbers of columns in **W** and **C**, respectively, as defined in Table I. However, at least one of (73a) and (74b) should be used in order to yield either $\mathbf{Q}_W$ or $\mathbf{P}_V$ which is required in calculating

$$\begin{aligned}(\mathbf{1} + \mathbf{W}^T\,\mathbf{A}^{-1}\,\mathbf{V}\,\mathbf{D}) &= (\mathbf{1} + {\mathbf{Q}_W}^T\,\mathbf{V}\,\mathbf{D}) \\ &= (\mathbf{1} + \mathbf{W}^T\,\mathbf{P}_V\,\mathbf{D})\,. \end{aligned} \tag{75}$$

Hence, according to the values of r_1, r_2, m' and n', we can choose appropriate approaches such that the computational effort is reduced. For example, when $m' < n'$ and $m' < r_2$, we can use

$$\Delta(\mathbf{B}^T\,\mathbf{A}^{-1}\,\mathbf{C}) = -\,\mathbf{S}^T\,{\mathbf{Q}_W}^T\,\mathbf{C}\,, \tag{76}$$

where **S** is the solution of the $r_2 \times r_2$ linear system

$$(1 + Q_W{}^T V D)^T S = RHS \qquad (77)$$

with the right-hand sides as

$$RHS = (Q_B{}^T V D)^T . \qquad (78)$$

This approach requires $m' + r_2$ FBS in the $n \times n$ linear system for Q_B and Q_W, one LU factorization and m' FBS in the $r_2 \times r_2$ linear system of (77).

As another example for computing (70), consider

$$\Delta(B^T A^{-1} C) = - B^T P_V D S , \qquad (79)$$

where S is the solution of the $r_2 \times r_2$ linear system

$$(1 + W^T P_V D) S = RHS \qquad (80)$$

with the right-hand sides as

$$RHS = W^T P_C . \qquad (81)$$

This approach requires $r_1 + n'$ FBS in the $n \times n$ linear system for P_V and P_C, one LU factorization and n' FBS in the $r_2 \times r_2$ linear system of (80).

Comparing the major computational effort required by the approaches of (76) and (79), we find that (76) can be recommended if $m' < n'$ and $r_2 + m' < r_1 + n'$, otherwise, if $m' > n'$ and $r_2 + m' > r_1 + n'$, (79) may be used.

Introduction to Tables VIII and IX

In Tables VIII and IX, we present an exhaustive search for all possible cases of $\Delta(B^T A^{-1} C)$ distinguished by the number of FBS in the $n \times n$ and $r_1 \times r_1$ or $r_2 \times r_2$ systems. Suitable expressions are formulated for efficient calculations of $\Delta(B^T A^{-1} C)$. Tables X and XI can be referred to for the number of LU factorizations and FBS required for each of these cases. We find that we always need one LU factorization in the original $n \times n$ matrix A, which is usually performed for the simulation of the original system before large changes and, another LU factorization in the $r_1 \times r_1$ or $r_2 \times r_2$ system, which is usually smaller than the original system and is performed only for the evaluation of large change effects. Different cases are classified such that the number of FBS in the $n \times n$ system equals the minimum of

$m' + r_2$, $n' + r_1$ and $r_1 + r_2$ and the number of FBS in the $r_1 \times r_1$ or $r_2 \times r_2$ smaller system equals the minimum of r_1, r_2, m′ and n′. This minimum FBS criterion can be used as a guide to select appropriate expressions among different cases of $\Delta(\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})$ as shown in Tables VIII and IX.

Also included in Tables VIII and IX, are formulas suitable for computations of large change effects of $\mathbf{A}^{-1}$ associated with vectors **b** and **c**.

Discussion and Duality Properties

Since (47) and (48) require the solution of linear systems of size $r_2 \times r_2$ and $r_1 \times r_1$, respectively, (47) is preferred for the case of $r_2 < r_1$ and (48) is preferred for $r_2 > r_1$. This is the phenomenon evident in Tables VIII and IX where two sets of formulas constitute a series of dual problems corresponding to $r_2 \leq r_1$ and $r_2 > r_1$, respectively. For example, consider the dual relations between

$$\Delta(\mathbf{b}^T \mathbf{A}^{-1} \mathbf{c}) = -\, \mathbf{s}^T \mathbf{Q}_W{}^T \mathbf{c}\,, \tag{82a}$$

where **s** is the solution of

$$(1 + \mathbf{Q}_W{}^T \mathbf{V}\, \mathbf{D})^T \mathbf{s} = (\mathbf{q}_b{}^T \mathbf{V}\, \mathbf{D})^T \tag{82b}$$

and

$$\Delta(\mathbf{b}^T \mathbf{A}^{-1} \mathbf{c}) = -\, \mathbf{b}^T \mathbf{P}_V\, \mathbf{s}, \tag{83a}$$

where **s** is the solution of

$$(1 + \mathbf{D}\, \mathbf{W}^T \mathbf{P}_V)\, \mathbf{s} = \mathbf{D}\, \mathbf{W}^T \mathbf{p}_c\,. \tag{83b}$$

If we apply the following interchanges

$$\mathbf{A} \leftrightarrow \mathbf{A}^T, \tag{84a}$$

$$\mathbf{b} \leftrightarrow \mathbf{c}\,, \tag{84b}$$

$$\mathbf{D} \leftrightarrow \mathbf{D}^T, \tag{84c}$$

$$\mathbf{V} \leftrightarrow \mathbf{W}, \tag{84d}$$

$$\mathbf{p}_c \leftrightarrow \mathbf{q}_b\,, \tag{84e}$$

and

$$\mathbf{P}_V \quad \leftrightarrow \quad \mathbf{Q}_W, \tag{84f}$$

then equations (82) and (83) are completely interchanged. In terms of computational effort, similar relations exist when we apply the interchanges

$$r_1 \quad \leftrightarrow \quad r_2 \tag{85a}$$

and

$$m' \quad \leftrightarrow \quad n'. \tag{85b}$$

For example, (82) requires $r_2 + 1$ FBS in the n x n linear system, one LU factorization and one FBS in an $r_2 \times r_2$ system. (83) requires $r_1 + 1$ FBS in the n x n system, one LU factorization and one FBS in an $r_1 \times r_1$ system.

Remarks

When the number of FBS exceeds the order of the system, we would rather perform a matrix inversion or, FBS operation no more than the order of the system. Together with necessary linear combination operations, the results required by all the FBS can be obtained.

Also notice that if **V** and **W** are chosen to contain variable parameters, other computational procedures can be formulated in order to use the large change formulas efficiently. However, it is recommended to arrange all variables contained in **D**. Thus, **V**, **W**, $\mathbf{P}_V$ and $\mathbf{Q}_W$ can be considered constant and saved for any changes of variable values. This feature can be found in Tables VIII and IX.

VIII. OPERATIONAL COUNTS FOR LARGE CHANGE FORMULAS WITH A^{-1} INVOLVED

Operational Counts

We all know that the number of product operations (multiplication or division) required for an n x n matrix LU factorization is $n^3/3 - n/3$, for a forward and backward substitution is n^2, and for a multiplication of an m x n matrix by an n x ℓ matrix is $m \cdot n \cdot \ell$.

Also, inverting an n x n matrix requires n^3 product operations [8,9]. Here, one quotient operation is considered equivalent to one product operation [2].

Consider the computation of $\Delta(\mathbf{A}^{-1})$. Suppose $r_1 + r_2 < n$ and the matrix $\mathbf{A}$ has already been LU factorized. In order to calculate $\Delta(\mathbf{A}^{-1})$ by (47), we need $(r_1 + r_2)\, n^2$ product (multiplication and division) operations to get $\mathbf{A}^{-1}\,\mathbf{V}$ and $\mathbf{W}^T\,\mathbf{A}^{-1}$ and $r_1 r_2\,(n + r_2)$ similar operations to obtain the r_2 x r_2 matrix $(\mathbf{1} + \mathbf{W}^T\,\mathbf{A}^{-1}\,\mathbf{V}\,\mathbf{D})$ which will be inverted after $r_2^{\,3}$ operations. Together with necessary matrix multiplications, the total operational count will be $n\,(nr_1 + nr_2 + r_1r_2) + r_2\,(2r_1r_2 + r_2^{\,2} + nr_1 + n^2)$.

If we use (64) to compute $\Delta(\mathbf{A}^{-1})$, we need to know the original inverse of $\mathbf{A}$. Suppose $\mathbf{A}^{-1}$ has been obtained. Then for each iteration, we perform $n(r_1 + 2 + 3n)$ operations. Altogether, we need $r_2\,n(r_1 + 2 + 3n)$ operations to obtain $\Delta(\mathbf{A}^{-1})$.

Suppose $r_1 > r_2$. We use Table VIII to calculate $\Delta(\mathbf{A}^{-1}\,\mathbf{c})$. We will need $r_1 n\,(n + r_2) + nr_2 + r_2^{\,3}/3 - r_2/3 + r_2^{\,2} + r_1 n + r_1 r_2\,(r_2 + 1)$ operations. If we use the iterative procedure of (64), the operational count will be $r_2 n\,(r_1 + 2 + 3n) + n^2$.

Discussion

If we can express $\Delta\mathbf{A}$ as $\mathbf{V}\,\mathbf{D}\,\mathbf{W}^T$ such that r_1 and r_2 are very small, using large change formulas can save computation, for example, up to 60% for $\Delta(\mathbf{A}^{-1})$ and 48% for $\Delta(\mathbf{A}^{-1}\,\mathbf{c})$, as compared with direct methods which calculate everything all over again.

If r_1 and r_2 are not small enough, the above computational counts of large change formulas will even exceed those of the direct methods. However, this computational count will be significantly reduced if variable parameters are changed many times. For example, in an optimization procedure, variables will usually be updated many times before the optimum is reached. Another example is centering and yield estimation, where Monte-Carlo procedures are often involved to choose random values repeatedly for variable parameters. When we apply large change formulas in these cases, we can choose $\mathbf{V}$, $\mathbf{D}$ and $\mathbf{W}$ such that $\mathbf{D}$ contains variables and $\mathbf{V}$ and $\mathbf{W}$ are constant. Thus all the FBS in the n x n system represented by **A**, i.e., $\mathbf{P}_V$ and $\mathbf{Q}_W$ are calculated only once and can be used for all subsequent changes of variables. It should be noted that the formulations of **V**, **D** and **W** in this case are designed only for a certain set of variables whose values will be changed more than once. For a different set of variables, **V**, **D** and **W** will be different hence $\mathbf{P}_V$ and $\mathbf{Q}_W$ have to be obtained by performing FBS or, under certain conditions, by performing linear combination operations. The operational count becomes, $r_2(2r_1r_2 + r_2^2 + nr_1 + n^2)$ for $\Delta(\mathbf{A}^{-1})$ and $r_2^3/3 - r_2/3 + r_2^2 + r_1n + r_1r_2(r_2 + 1)$ for $\Delta(\mathbf{A}^{-1}\mathbf{c})$ using formulas corresponding to (47), for large changes of the second time and on.

Tables XII and XIII show the operational counts of calculating $\Delta(\mathbf{A}^{-1})$ and $\Delta(\mathbf{A}^{-1}\mathbf{c})$ by using the direct method and by using our large change formulas. The direct method refers to the method of calculating the system all over again after a large change. Computer programs calculating $\Delta(\mathbf{A}^{-1})$ by formula (64) and $\Delta(\mathbf{A}^{-1}\mathbf{c})$ by (69) have been developed which confirm the operational count shown in Tables XII and XIII.

From Tables XII and XIII, we can see that when r_1 and r_2 are small, all large change formulas are very efficient. As r_1 and r_2 increase, the scheme of (64) tends to lose efficiency whereas formula (47) (similarly (48), (55), (56)) can be still preferred if variables are changed more than once.

IX. ON THE FORMULATION OF V, D and W

Introduction

From the previous discussion, we find that the calculation of large change effects with $\mathbf{A}^{-1}$ involved is essentially reduced to solving a smaller system of size r_1 by r_1 or r_2 by r_2, where r_1 and r_2 are the numbers of rows and columns in **D**, respectively. If $r_1 \geq r_2$, we use (47) for an $r_2 \times r_2$ smaller system, otherwise use (48) yielding an $r_1 \times r_1$ system. Different formulations of **V**, **W** and **D** can lead to different computational cost which depends heavily upon the value of r_1 (if $r_1 < r_2$) or r_2 (if $r_1 > r_2$). Thus, it is always desirable to have the value of either r_1 or r_2 as low as possible. According to (13), we can always choose, without computation,

$$\mathbf{V} = \mathbf{U}_I , \tag{86}$$

$$\mathbf{W} = \mathbf{U}_J \tag{87}$$

and

$$\mathbf{D} = \Delta\mathbf{A}_{IJ} . \tag{88}$$

In this way, $r_1 = n_I$, $r_2 = n_J$. Therefore, $r_1 \leq n$ and $r_2 \leq n$. So, we are assured that this formulation results in solving a system no greater than the original n x n system.

Usually, r_1 or r_2 can be further reduced. It can be shown that the lower bound for r_1 or r_2 exists such that **D** is of dimension $r_1 \times r_2$ and (11) holds.

<u>The Minimum Value for r_1 and r_2</u>

Evidently,

$$\begin{aligned} \text{Rank}\,(\Delta\mathbf{A}) = &\ \text{Rank}\,(\mathbf{V}\,\mathbf{D}\,\mathbf{W}^T) \\ \leq &\ \min\,\{\text{Rank}\,\mathbf{V}, \text{Rank}\,\mathbf{D}, \text{Rank}\,\mathbf{W}\} \\ \leq &\ \min\,\{n, r_1, r_2\} . \end{aligned} \tag{89}$$

Let

$$r \triangleq \text{Rank}\,(\Delta\mathbf{A}) . \tag{90}$$

Then r is a lower bound for r_1 and r_2.

Also, we can always choose **V**, **W** and **D** such that $r_1 = r$ or $r_2 = r$, i.e., the lower bound of r_1 and r_2 can always be reached. For example, since r = Rank ($\Delta\mathbf{A}$), $\Delta\mathbf{A}$ has r linearly independent columns, the linear combination of which gives all columns of $\Delta\mathbf{A}$. We can choose **V** as

$$\mathbf{V} = [\mathbf{v}_1\ \mathbf{v}_2\ \ldots\ \mathbf{v}_r] \tag{91}$$

such that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are the r independent columns of $\Delta\mathbf{A}$. $\Delta\mathbf{A}$ can be expressed by linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ as

$$\Delta \mathbf{A} = \left[\sum_{i=1}^{r} (w_{1i}\, \mathbf{v}_i) \;\; \sum_{i=1}^{r} (w_{2i}\, \mathbf{v}_i) \;\; \cdots \;\; \sum_{i=1}^{r} (w_{ni}\, \mathbf{v}_i) \right]. \tag{92}$$

Let

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1r} \\ w_{21} & w_{22} & \cdots & w_{2r} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ w_{n1} & w_{n2} & \cdots & w_{nr} \end{bmatrix}. \tag{93}$$

Then we have

$$\Delta \mathbf{A} = \mathbf{V}\,\mathbf{W}^T = \mathbf{V}\,\mathbf{D}\,\mathbf{W}^T, \tag{94}$$

where $\mathbf{D}$ is an r x r identity matrix and $r_1 = r_2 = r$.

Therefore, we have

$$\min_{(\mathbf{V},\mathbf{D},\mathbf{W})} r_1 = \min_{(\mathbf{V},\mathbf{D},\mathbf{W})} r_2 = r\,. \tag{95}$$

The above equation gives the conclusion that for evaluating large changes involving $\mathbf{A}^{-1}$, the minimum size of the system we need to solve is r by r, where r is the rank of $\Delta\mathbf{A}$.

Discussion

In persuing the minimum size system, additional computations will probably be involved to construct **V**, **D** and **W** before using large change formulas. The objective to decrease r_1 or r_2 is to reduce computations in evaluating large change effects. Therefore, there is a trade-off between reducing r_1 or r_2 and the computations caused by this reduction. The objective should be such that the overall large change computation is reduced. In practice, one can start from the formulation of (13), which requires no additional computation, and then persuing the reduction of r_1 or r_2 if only little effort needs to be involved. The actual calculations are usually performed with $r \le r_1 \le n_I$ and $r \le r_2 \le n_J$.

However, in practical problems, e.g., in computer-aided circuit design, the physical properties of the problem can lead to a formulation of **V**, **W** and **D** such that the minimum size system is achieved and no additional computations are introduced.

X. NUMERICAL EXAMPLES

Example 1

To show how our first-order and large change sensitivity formulas are used, we consider, as an example, a set of linear equations,

$$\mathbf{A}\,\mathbf{x} = \mathbf{b}\,, \tag{96}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 0 \\ 3 & 1 & 4 & 1 \\ 6 & 2 & 1 & -1 \end{bmatrix} \tag{97}$$

and

$$\mathbf{b} = [11 \quad 8 \quad 23 \quad 13]^T\,. \tag{98}$$

The solution can be easily found as

$$\mathbf{x} = [\,2 \quad 1 \quad 3 \quad 4]^T\,. \tag{99}$$

Suppose elements of **A** at rows 1, 2 and 4 and columns 2 and 3 are changed yielding $\mathbf{A}_{new}$ such that

$$\mathbf{A}_{new} = \mathbf{A} + \Delta\mathbf{A}\,, \tag{100}$$

where

$$\Delta\mathbf{A} = \begin{bmatrix} 0 & -2 & 1 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 \end{bmatrix}. \tag{101}$$

Instead of solving

$$\mathbf{A}_{new}\,\mathbf{x}_{new} = \mathbf{b} \tag{102}$$

for $\mathbf{x}_{new}$ directly, we can use large change formulas. In this example, we choose such a small system in order to facilitate the demonstration and to guide through the use of formulas easily step by step.

Evidently, the sets I and J, defined in (4) and (5), are

$$I = \{1, 2, 4\} \tag{103}$$

and

$$J = \{2, 3\}. \tag{104}$$

Therefore, $\mathbf{U}_I$ and $\mathbf{U}_J$ defined in (6) and (7), are

$$\mathbf{U}_I = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_4] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{105}$$

and

$$\mathbf{U}_J = [\mathbf{u}_2 \ \mathbf{u}_3] = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{106}$$

Also,

$$\Delta\mathbf{A}_{IJ} = \begin{bmatrix} -2 & 1 \\ 3 & 0 \\ -1 & 2 \end{bmatrix}. \tag{107}$$

Let $\mathbf{V} = \mathbf{U}_I$, $\mathbf{W} = \mathbf{U}_J$ and $\mathbf{D} = \Delta\mathbf{A}_{IJ}$, then we have successfully formulated $\mathbf{V}$, $\mathbf{W}$ and $\mathbf{D}$ such that (11) holds. Formally, we have

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \tag{108}$$

Applying large change operator Δ to both sides and using (47), since $r_1 = 3 > 2 = r_2$, we obtain,

$$\begin{aligned}\Delta \mathbf{x} = &\ \Delta(\mathbf{A}^{-1}\mathbf{b})\\ = &\ -\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D}(1+\mathbf{W}^T\mathbf{A}^{-1}\mathbf{V}\,\mathbf{D})^{-1}\mathbf{W}^T\mathbf{A}^{-1}\mathbf{b}\\ = &\ -\mathbf{P}_V\,\mathbf{D}(1+\mathbf{W}^T\mathbf{P}_V\,\mathbf{D})^{-1}\mathbf{W}^T\mathbf{x}\\ = &\ -\mathbf{P}_V\,\mathbf{D}\,\mathbf{s}\,,\end{aligned} \tag{109}$$

where $\mathbf{s}$ is similarly defined in calculating $\Delta(\mathbf{A}^{-1}\mathbf{c})$ in Table VIII corresponding to $r_1 \geq r_2$. $\mathbf{P}_V$ is obtained by performing 3 FBS with the R.H.S. as $\mathbf{V}$, i.e.,

$$\mathbf{A}\,\mathbf{P}_V = \mathbf{V}\,, \tag{110}$$

where $\mathbf{A}$ has already been LU factorized when we solve (96) for $\mathbf{x}$ of (99) and where $\mathbf{V} = \mathbf{U}_I$. The solution of $\mathbf{P}_V$ is

$$\mathbf{P}_V = \begin{bmatrix} 0.50 & -3.50 & 1.00 \\ -0.25 & 4.25 & -1.00 \\ -0.75 & 3.75 & -1.00 \\ 1.75 & -8.75 & 2.00 \end{bmatrix} \tag{111}$$

and the solution of

$$(1+\mathbf{W}^T\mathbf{P}_V\,\mathbf{D})\,\mathbf{s} = \mathbf{W}^T\mathbf{x}\,, \tag{112}$$

i.e.,

$$\begin{bmatrix} 15.25 & -2.25 \\ 13.75 & -1.75 \end{bmatrix} \cdot \mathbf{s} = \begin{bmatrix} 1.00 \\ 3.00 \end{bmatrix} \tag{113}$$

is

$$\mathbf{s} = \begin{bmatrix} 1.17647 \\ 7.52941 \end{bmatrix} . \tag{114}$$

Therefore,

$$\begin{aligned}\Delta \mathbf{x} = &\ -\mathbf{P}_V\ \mathbf{D}\,\mathbf{s}\\ = &\begin{bmatrix} -4.11765 \\ 0.17647 \\ 4.52941 \\ -5.94118 \end{bmatrix}\end{aligned} \tag{115}$$

and the solution of (102) can be found by

$$\mathbf{x}_{new} = \mathbf{x} + \Delta\mathbf{x}$$

$$= \begin{bmatrix} -2.11765 \\ 1.17647 \\ 7.52941 \\ -1.94118 \end{bmatrix}. \qquad (116)$$

If the intersection elements of **A** in rows 1, 2, 4 and columns 2, 3 change again, yielding $\mathbf{A}'_{new}$,

$$\mathbf{A}'_{new} = \mathbf{A} + \begin{bmatrix} 0 & -1.5 & 1.5 & 0 \\ 0 & 1.5 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2.2 & 4 & 0 \end{bmatrix}, \qquad (117)$$

the solution of

$$\mathbf{A}'_{new}\, \mathbf{x}'_{new} = \mathbf{b} \qquad (118)$$

can be obtained using not only the LU factorized **A**, but also $\mathbf{P}_V$, $\mathbf{W}^T\mathbf{x}$ and $\mathbf{W}^T\, \mathbf{P}_V$ which we already calculated. The computation is then to formulate and to solve the 2 x 2 system

$$(\mathbf{1} + \mathbf{W}^T\, \mathbf{P}_V\, \mathbf{D}')\, \mathbf{s}' = \mathbf{W}^T\, \mathbf{x} \qquad (119)$$

to obtain **s**′, where

$$\mathbf{D}' = \Delta\mathbf{A}_{IJ} = \begin{bmatrix} -1.5 & 1.5 \\ 1.5 & 1 \\ -2.2 & 4 \end{bmatrix} \qquad (120)$$

and where

$$\mathbf{s}' = \begin{bmatrix} 0 \\ -8 \end{bmatrix}. \qquad (121)$$

The solution of (118) is obtained, by using (109) as

$$\mathbf{x}'_{new} = \mathbf{x} + \Delta\mathbf{x} = \mathbf{x} - \mathbf{P}_V\, \mathbf{D}'\, \mathbf{s}'$$

$$= [12 \;\; 0 \;\; -8 \;\; 19]^T. \qquad (122)$$

For further changes of element of **A** at rows i, $i \in I$, and columns j, $j \in J$, we repeat the procedure from (119) to (122) to obtain corresponding solutions.

In this example, $n = 4$, $r_1 = n_I = 3$ and $r_2 = n_J = 2$. This example is used only for demonstration since directly solving (102) and (118) is more economical. Large change formulas can increase efficiency only when we have a large system with relatively small r_1 or r_2, as shown in Table XIII and in Example 3.

Example 2

To show the different formulations of **V**, **D** and **W**, we again consider (96) but

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 3 & 0 & 7 & 2 \\ 2 & 3 & 1 & 2 & 4 & 5 & 2 \\ 3 & 4 & 0 & 2 & 1 & 2 & 4 \\ 6 & 5 & 0 & 2 & 7 & 1 & 3 \\ 8 & 3 & 2 & 6 & 4 & 4 & 9 \\ 4 & 0 & 2 & 6 & 5 & 5 & 2 \\ 7 & 1 & 5 & 4 & 2 & 0 & 5 \end{bmatrix} \qquad (123)$$

and

$$\mathbf{b} = [76 \quad 83 \quad 64 \quad 86 \quad 151 \quad 103 \quad 85]^T . \qquad (124)$$

We find the solution of the system is

$$\mathbf{x} = [1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7]^T . \qquad (125)$$

Suppose elements of **A** at rows i, $i \in I$, and columns j, $j \in J$, are changed, where

$$I = \{1, 3, 4, 6, 7\}, \qquad (126)$$

$$J = \{2, 4, 5, 7\}, \qquad (127)$$

and

$$\Delta \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 3 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 1 \\ 0 & 2 & 0 & 3 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 3 & 0 & 4 & 2 & 0 & 1 \end{bmatrix} . \tag{128}$$

The following shows three different formulations of **V**, **D** and **W** such that (11) is satisfied. The easiest formulation is to choose $\mathbf{V} = \mathbf{U}_I$, $\mathbf{W} = \mathbf{U}_J$ and $\mathbf{D} = \Delta\mathbf{A}_{IJ}$. In this example,

$$\mathbf{U}_I = [\mathbf{u}_1 \;\; \mathbf{u}_3 \;\; \mathbf{u}_4 \;\; \mathbf{u}_6 \;\; \mathbf{u}_7] , \tag{129}$$

$$\mathbf{U}_J = [\mathbf{u}_2 \;\; \mathbf{u}_4 \;\; \mathbf{u}_5 \;\; \mathbf{u}_7] , \tag{130}$$

and

$$\Delta \mathbf{A}_{IJ} = \begin{bmatrix} 1 & 3 & -1 & 2 \\ 0 & 1 & -1 & 1 \\ 2 & 3 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 3 & 4 & 2 & 1 \end{bmatrix} , \tag{131}$$

where the subscripted vectors **u** are unit 7-vectors defined in Section II. This formulation requires 5 FBS to obtain $\mathbf{P}_V$ and the solution of a 4 by 4 system corresponding to (112). Here $r_1 = 5$ and $r_2 = 4$.

Notice that the first column of $\Delta\mathbf{A}_{IJ}$ can be obtained by adding the third and fourth columns. Then r_2 is reduced to 3 since we can express

$$\mathbf{V} = \mathbf{U}_I , \tag{132}$$

$$\mathbf{D} = \begin{bmatrix} 3 & -1 & 2 \\ 1 & -1 & 1 \\ 3 & 1 & 1 \\ 1 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix} \tag{133}$$

and

$$\mathbf{W} = [\mathbf{u}_4 \quad \mathbf{u}_5+\mathbf{u}_2 \quad \mathbf{u}_7+\mathbf{u}_2]$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{134}$$

Thus the system of (112) becomes 3 by 3.

Further reduction of r_2 can be achieved by considering the first column of (133) as the linear combination of other columns. Thus, we can choose

$$\mathbf{V} = \mathbf{U}_I , \tag{135}$$

$$\mathbf{D} = \begin{bmatrix} -1 & 2 \\ -1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 2 & 1 \end{bmatrix} \tag{136}$$

and

$$\mathbf{W} = [\mathbf{u}_2 + \mathbf{u}_4 + \mathbf{u}_5 \quad \mathbf{u}_2 + 2\mathbf{u}_4 + \mathbf{u}_7]$$

$$= \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ 1 & 2 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} . \tag{137}$$

Notice that here $r_2 = 2$ and no further reduction in r_2 can be obtained since

$$\text{Rank}(\Delta\mathbf{A}) = \text{Rank}(\Delta\mathbf{A}_{IJ}) = 2 . \tag{138}$$

r_1 can also be reduced to 2 by defining **V** and **D** as

$$\mathbf{V} = [2\mathbf{u}_1 + \mathbf{u}_3 + \mathbf{u}_4 + \mathbf{u}_7 \quad -3\mathbf{u}_1 - 2\mathbf{u}_3 + \mathbf{u}_6 + \mathbf{u}_7]$$

$$= \begin{bmatrix} 2 & -3 \\ 0 & 0 \\ 1 & -2 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} , \tag{139}$$

$$\mathbf{D} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \tag{140}$$

and by choosing **W** the same as (137). By this formulation of **V**, **D** and **W**, we only need to perform 2 FBS to obtain $\mathbf{P}_V$ and to solve a 2 by 2 system corresponding to (112).

Using the last formulation of **V**, **D** and **W**, $\mathbf{P}_V$ is obtained, after performing forward and backward substitutions indicated by (110), as

$$\mathbf{P}_V = \begin{bmatrix} 1.17484 & -0.84026 \\ 0.25571 & -0.46405 \\ -0.26520 & 0.54841 \\ -0.15341 & -0.03053 \\ -0.70369 & 0.86208 \\ 0.38460 & -0.41274 \\ -0.82651 & 0.60036 \end{bmatrix} . \tag{141}$$

The 2 by 2 system is formulated by (112) which gives

$$\begin{bmatrix} 0.76610 & -0.60139 \\ -0.80238 & 0.12237 \end{bmatrix} \mathbf{s} = \begin{bmatrix} 11.00000 \\ 17.00000 \end{bmatrix} . \tag{142}$$

Solving this small system, we have

$$\mathbf{s} = \begin{bmatrix} -29.75797 \\ -56.19887 \end{bmatrix} . \tag{143}$$

The final solution is then obtained as

$$\mathbf{x}_{new} = \mathbf{x} + \Delta\mathbf{x} = \mathbf{x} - \mathbf{P}_V \, \mathbf{D} \, \mathbf{s}$$

$$= \begin{bmatrix} 76.98124 \\ 10.17098 \\ -3.47626 \\ -10.09545 \\ -29.83350 \\ 26.77673 \\ -46.17895 \end{bmatrix} . \tag{144}$$

If we directly solve (102) by performing a new LU factorization and forward and backward substitution, we obtain the same result as (144).

Example 3

The problem of solving a system of linear equations and evaluating the effects in the solution due to changes in the coefficient matrix is frequently encountered in practice. A computer program has been developed to implement the method used by the previous two examples. We formulate a 10 by 10 system, where the elements at the intersections of the 2nd, 5th, 9th rows and 3rd and 6th columns are constantly changed. We formulate **V**, **D** and **W** by (13). **V**, **W**, $\mathbf{P}_V$, $\mathbf{W}^T\mathbf{P}_V$ and $\mathbf{W}^T\mathbf{x}$ are calculated only once. For different values of variable elements, the corresponding **D** is different and we only need to formulate and to solve the 2 by 2 (r_2 by r_2) system of (112) to obtain **s** and calculate the new solution $\mathbf{x}_{new}$ by using

$$\mathbf{x}_{new} = \mathbf{x} + \Delta\mathbf{x} = \mathbf{x} - \mathbf{P}_V\,\mathbf{D}\,\mathbf{s}\,. \tag{145}$$

Numerical solutions as well as intermediate results are shown in Fig. 1. Fig. 2 shows the efficiency of this method compared with the direct method which is to solve the n by n (here 10 by 10) system repeatedly. For example, if the variables change 5 times, then the computational cost can be reduced by 60%, if they change 20 times, then 80%. Notice that in Fig. 2, the horizontal axis represents the number of simulations which include the initial system simulation and subsequent simulations when the system parameters are changed.

XI. CONCLUSIONS

We have presented a unified treatment of exact first-order and large change sensitivity problems in linear systems. Computational aspects of this approach have been discussed. Besides their computational efficiency, the sensitivity formulas also give us further insight into the relations between system parameters and system outputs, thus helping us in systems analysis and design. For a system with almost all of its parameters under change, we do not recommend the use of large change formulas. However, if variables exist only in some local areas of the system, using our large change formulas can save a considerable amount of computation. This philosophy can be compared to the use of sparse techniques where non-zero elements represents only a small number of the whole system elements. Besides directly applicable to linear systems, our formulas can also be used for gradient calculations and in nonlinear systems involving solutions of linear equations. Relevant expressions presented in the text and shown in tables can provide a useful tool for engineering designers to yield fast and efficient design procedures.

REFERENCES

[1] J.C.G. Boot, Quadratic Programming: Algorithms - Anomalies -Applications. Amsterdam: North-Holland, 1964, Chap. 2.

[2] A.S. Householder, Principles of Numerical Analysis. New York: McGraw- Hill, 1953, Chap. 2.

[3] A.S. Householder, "A survey of some closed methods for inverting matrices", SIAM J., vol. 5, 1957, pp. 155-169.

[4] J. Sherman and W.J. Morrison, "Adjustment of an inverse matrix corresponding to changes in the elements of a given column or a given row of the original matrix", Ann. Math. Stat., vol. 20, 1949, p. 621.

[5] J. Sherman and W.J. Morrison, "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix", Ann. Math. Stat., vol. 21, 1950, pp. 124-127.

[6] J.K. Reid, Ed., Large Sparse Sets of Linear Equations. London: Academic Press, 1971, pp. 41-56.

[7] S. Gupta, "An algorithm to compute large-change sensitivities of linear digital networks", Proc. IEEE, vol. 67, 1979, pp. 954-956.

[8] P.W. Williams, Numerical Computation. London: Nelson, 1972, Chap. 4.

[9] L. Fox, An Introduction to Numerical Linear Algebra. Oxford: Clarendon Press, 1964, Chap. 7.

TABLE I

BASIC NOTATION

Category	Notation	Dimension	Comment
Matrices	$\mathbf{A}$	m x n	containing variables $\boldsymbol{\phi}$. m = n if $\mathbf{A}^{-1}$ is involved.
	$\mathbf{A}_{IJ}$	n_I x n_J	submatrix of **A**.
	$\mathbf{B}$	m x m′	constant.
	$\mathbf{C}$	n x n′	constant.
	$\mathbf{D}$	r_1 x r_2	such that $\Delta\mathbf{A} = \mathbf{V}\,\mathbf{D}\,\mathbf{W}^T$.
	$\mathbf{P}$	n x n_{RHS}	* solution of $\mathbf{AP} = \mathbf{RHS}$.
	$\mathbf{Q}$	n x n_{RHS}	* solution of $\mathbf{A}^T\mathbf{Q} = \mathbf{RHS}$.
	$\mathbf{U}_I$	n x n_I	containing unit vectors.
	$\mathbf{U}_J$	n x n_J	containing unit vectors.
	$\mathbf{V}$	m x r_1	such that $\Delta\mathbf{A} = \mathbf{V}\,\mathbf{D}\,\mathbf{W}^T$.
	$\mathbf{W}$	n x r_2	such that $\Delta\mathbf{A} = \mathbf{V}\,\mathbf{D}\,\mathbf{W}^T$.

TABLE I (continued)

Category	Notation	Dimension	Comment
Vectors	$\mathbf{b}$	m	constant.
	$\mathbf{c}$	n	constant.
	$\mathbf{p}$	n	* solution of $\mathbf{Ap} = \mathbf{RHS}$.
	$\mathbf{q}$	n	* solution of $\mathbf{A}^T\mathbf{q} = \mathbf{RHS}$.
	$\mathbf{u}_i$	m	unit vectors of appropriate dimensions such that the premultiplications and postmultiplications in corresponding formulas exist.
	$\mathbf{u}_j$	n	
	$\mathbf{u}_\ell$	m′	
	$\mathbf{u}_k$	n′	

* Various subscripts will be added to the notation for various **RHS**. Table II can be referred to for the relations of these subscripts, **RHS** and n_{RHS}.

TABLE II

DEFINITIONS OF MATRICES **P** AND **Q** AND VECTORS **p** AND **q**

ASSOCIATED WITH VARIOUS SUBSCRIPTS

(* denotes **P** or **p**, † denotes **Q** or **q**)

(*) or (†)	Subscript of (*) or (†)	R.H.S. of $\mathbf{A}(*) = \mathbf{RHS}$ or $\mathbf{A}^T(\dagger) = \mathbf{RHS}$	Dimension of (*) or (†)
P, **Q**	B	**B**	n x m′
	C	**C**	n x n′
	V	**V**	$n \times r_1$
	W	**W**	$n \times r_2$
	U_I	$\mathbf{U}_I$	$n \times n_I$
	U_J	$\mathbf{U}_J$	$n \times n_J$
p, **q**	b	**b**	n
	c	**c**	n
	u_i	$\mathbf{u}_i$	n
	u_j	$\mathbf{u}_j$	n

TABLE III

FIRST-ORDER SENSITIVITIES W.R.T. COMPONENTS OF A MATRIX OR A VECTOR

Identification	Sensitivity Expression	
$\dfrac{\partial \mathbf{A}}{\partial A_{ij}}$	$\mathbf{u}_i \mathbf{u}_j^T$	
$\dfrac{\partial(\mathbf{b}^T \mathbf{A})}{\partial \mathbf{b}}$	$\mathbf{A}$	
$\dfrac{\partial(\mathbf{A}\,\mathbf{c})^T}{\partial \mathbf{c}}$	$\mathbf{A}^T$	
$\dfrac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{c})}{\partial \mathbf{A}}$	$\mathbf{b}\,\mathbf{c}^T$	
$\dfrac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{c})}{\partial \mathbf{b}}$	$\mathbf{A}\,\mathbf{c}$	
$\dfrac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{c})}{\partial \mathbf{c}}$	$\mathbf{A}^T \mathbf{b}$	
$\dfrac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{b})}{\partial \mathbf{b}}$	$(\mathbf{A} + \mathbf{A}^T)\mathbf{b}$	†
$\dfrac{\partial(\mathbf{B}^T \mathbf{A}\,\mathbf{C})}{\partial A_{ij}}$	$\mathbf{B}^T \mathbf{u}_i \mathbf{u}_j^T \mathbf{C}$	
$\dfrac{\partial[\mathbf{B}^T \mathbf{A}\,\mathbf{C}]_{\ell k}}{\partial \mathbf{A}}$	$\mathbf{B}\,\mathbf{u}_\ell \mathbf{u}_k^T \mathbf{C}^T$	

TABLE III (continued)

Identification	Sensitivity Expression
$\frac{\partial \mathbf{A}^{-1}}{\partial A_{ij}}$	$-\mathbf{A}^{-1}\mathbf{u}_i\mathbf{u}_j^T\mathbf{A}^{-1}$
$\frac{\partial(\mathbf{b}^T\mathbf{A}^{-1}\mathbf{c})}{\partial \mathbf{A}}$	$-(\mathbf{A}^{-1})^T\mathbf{b}\,\mathbf{c}^T(\mathbf{A}^{-1})^T$
$\frac{\partial(\mathbf{B}^T\mathbf{A}^{-1}\mathbf{C})}{\partial A_{ij}}$	$-\mathbf{B}^T\mathbf{A}^{-1}\mathbf{u}_i\mathbf{u}_j^T\mathbf{A}^{-1}\mathbf{C}$
$\frac{\partial[\mathbf{B}^T\mathbf{A}^{-1}\mathbf{C}]_{\ell k}}{\partial \mathbf{A}}$	$-(\mathbf{B}^T\mathbf{A}^{-1})^T\mathbf{u}_\ell\mathbf{u}_k^T(\mathbf{A}^{-1}\mathbf{C})^T$

† In this case, **A** is a square matrix (m = n).

TABLE IV

FIRST-ORDER SENSITIVITY W.R.T. COMPONENTS OF MATRIX **A** WHEN **A** IS SYMMETRICAL AND $i \neq j$

Identification	Sensitivity Expression ($\mathbf{A} = \mathbf{A}^T$)
$\frac{\partial \mathbf{A}}{\partial A_{ij}}$	$\mathbf{u}_i \mathbf{u}_j^T + \mathbf{u}_j \mathbf{u}_i^T$
$\frac{\partial(\mathbf{b}^T \mathbf{A})}{\partial \mathbf{b}}$	$\mathbf{A}$
$\frac{\partial(\mathbf{A}\,\mathbf{c})^T}{\partial \mathbf{c}}$	$\mathbf{A}$
$\frac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{c})}{\partial \mathbf{A}}$	$\mathbf{b}\,\mathbf{c}^T + \mathbf{c}\,\mathbf{b}^T$
$\frac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{c})}{\partial \mathbf{b}}$	$\mathbf{A}\,\mathbf{c}$
$\frac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{c})}{\partial \mathbf{c}}$	$\mathbf{A}\,\mathbf{b}$
$\frac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{b})}{\partial \mathbf{b}}$	$2\mathbf{A}\,\mathbf{b}$
$\frac{\partial(\mathbf{B}^T \mathbf{A}\,\mathbf{C})}{\partial A_{ij}}$	$\mathbf{B}^T(\mathbf{u}_i \mathbf{u}_j^T + \mathbf{u}_j \mathbf{u}_i^T)\mathbf{C}$
$\frac{\partial[\mathbf{B}^T \mathbf{A}\,\mathbf{C}]_{\ell k}}{\partial \mathbf{A}}$	$\mathbf{B}\,\mathbf{u}_\ell \mathbf{u}_k^T \mathbf{C}^T + \mathbf{C}\,\mathbf{u}_k \mathbf{u}_\ell^T \mathbf{B}^T$

TABLE IV (continued)

Identification	Sensitivity Expression ($\mathbf{A} = \mathbf{A}^T$)
$\dfrac{\partial \mathbf{A}^{-1}}{\partial A_{ij}}$	$-\mathbf{A}^{-1}(\mathbf{u}_i \mathbf{u}_j{}^T + \mathbf{u}_j \mathbf{u}_i{}^T)\mathbf{A}^{-1}$
$\dfrac{\partial (\mathbf{b}^T \mathbf{A}^{-1} \mathbf{c})}{\partial \mathbf{A}}$	$-\mathbf{A}^{-1}(\mathbf{b}\,\mathbf{c}^T + \mathbf{c}\,\mathbf{b}^T)\mathbf{A}^{-1}$
$\dfrac{\partial (\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})}{\partial A_{ij}}$	$-\mathbf{B}^T \mathbf{A}^{-1}(\mathbf{u}_i \mathbf{u}_j{}^T + \mathbf{u}_j \mathbf{u}_i{}^T)\mathbf{A}^{-1} \mathbf{C}$
$\dfrac{\partial [\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}]_{\ell k}}{\partial \mathbf{A}}$	$-\mathbf{A}^{-1}(\mathbf{B}\,\mathbf{u}_\ell \mathbf{u}_k{}^T \mathbf{C}^T + \mathbf{C}\,\mathbf{u}_k \mathbf{u}_\ell{}^T \mathbf{B}^T)\mathbf{A}^{-1}$

TABLE V

EXPRESSIONS APPROPRIATE FOR COMPUTATIONS FOR SENSITIVITIES W.R.T. COMPONENTS OF MATRIX **A** WHEN $\mathbf{A}^{-1}$ IS INVOLVED

Identification	Sensitivity Expression (a) General	Sensitivity Expression (b) when $\mathbf{A} = \mathbf{A}^T$ and $i \neq j$
$\dfrac{\partial \mathbf{A}^{-1}}{\partial A_{ij}}$	$-\mathbf{p}_{ui}\,\mathbf{q}_{uj}^T$	$-(\mathbf{p}_{ui}\,\mathbf{p}_{uj}^T + \mathbf{p}_{uj}\,\mathbf{p}_{ui}^T)$
$\dfrac{\partial(\mathbf{b}^T \mathbf{A}^{-1} \mathbf{c})}{\partial \mathbf{A}}$	$-\mathbf{q}_b\,\mathbf{p}_c^T$	$-(\mathbf{p}_b\,\mathbf{p}_c^T + \mathbf{p}_c\,\mathbf{p}_b^T)$
$\dfrac{\partial(\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})}{\partial A_{ij}}$	$-\mathbf{B}^T\,\mathbf{p}_{ui}\,\mathbf{q}_{uj}^T\,\mathbf{C}$	$-\mathbf{B}^T(\mathbf{p}_{ui}\,\mathbf{p}_{uj}^T + \mathbf{p}_{uj}\,\mathbf{p}_{ui}^T)\mathbf{C}$
$\dfrac{\partial[\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C}]_{\ell k}}{\partial \mathbf{A}}$	$-\mathbf{q}_b\,\mathbf{p}_c^T$	$-(\mathbf{p}_b\,\mathbf{p}_c^T + \mathbf{p}_c\,\mathbf{p}_b^T)$ †

† where **b** is the ℓth column of **B** and **c** is the kth column of **C**. Both **b** and **c** are used as the R.H.S. of the system involving **A** for original solutions $\mathbf{p}_b$, $\mathbf{p}_c$ and adjoint solution $\mathbf{q}_b$.

TABLE VI

EXPRESSIONS APPROPRIATE FOR COMPUTATION OF SENSITIVITIES

W.R.T. VARIABLE ϕ WHEN $\mathbf{A}^{-1}$ IS INVOLVED

Identification	Sensitivity Expression
$\dfrac{\partial \mathbf{A}^{-1}}{\partial \phi}$	$-\mathbf{P}_{UI} \dfrac{\partial \mathbf{A}_{IJ}}{\partial \phi} \mathbf{Q}_{UJ}^T$
$\dfrac{\partial (\mathbf{b}^T \mathbf{A}^{-1})}{\partial \phi}$	$-(\mathbf{q}_b^T)_I \dfrac{\partial \mathbf{A}_{IJ}}{\partial \phi} \mathbf{Q}_{UJ}^T$
$\dfrac{\partial (\mathbf{A}^{-1} \mathbf{c})}{\partial \phi}$	$-\mathbf{P}_{UI} \dfrac{\partial \mathbf{A}_{IJ}}{\partial \phi} (\mathbf{p}_c)_J$
$\dfrac{\partial (\mathbf{b}^T \mathbf{A}^{-1} \mathbf{c})}{\partial \phi}$	$-(\mathbf{q}_b^T)_I \dfrac{\partial \mathbf{A}_{IJ}}{\partial \phi} (\mathbf{p}_c)_J$
$\dfrac{\partial (\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})}{\partial \phi}$	$-(*) \dfrac{\partial \mathbf{A}_{IJ}}{\partial \phi} (\dagger)$

$$(*) = \begin{cases} \mathbf{B}^T \mathbf{P}_{UI} & \text{if } n_I < m' \\ \mathbf{Q}_B^T \mathbf{U}_I & \text{if } n_I \geq m' \end{cases}, \qquad (\dagger) = \begin{cases} \mathbf{Q}_{UJ}^T \mathbf{C} & \text{if } n_J < n' \\ \mathbf{U}_J^T \mathbf{P}_C & \text{if } n_J \geq n' \end{cases}$$

$(\mathbf{q}_b)_I$ and $(\mathbf{p}_c)_J$ are defined as vectors consisting of all ith elements of $\mathbf{q}_b$, $i \in I$, and all jth elements of $\mathbf{p}_c$, $j \in J$, respectively.

TABLE VII

LARGE CHANGE FORMULAS FOR THE INVERSE OF A MATRIX

Formulation of $\Delta\mathbf{A}$ as $\mathbf{V\,D\,W}^T$		Formulas for $\Delta(\mathbf{A}^{-1})$
$\mathbf{V\,D\,W}^T$		$-\ \mathbf{A}^{-1}\mathbf{V\,D}(1+\mathbf{W}^T\mathbf{A}^{-1}\mathbf{V\,D})^{-1}\mathbf{W}^T\mathbf{A}^{-1}$
(general form)		$-\ \mathbf{A}^{-1}\mathbf{V}(1+\mathbf{D\,W}^T\mathbf{A}^{-1}\mathbf{V})^{-1}\mathbf{D\,W}^T\mathbf{A}^{-1}$
	†	$-\ \mathbf{A}^{-1}\mathbf{V}(\mathbf{D}^{-1}+\mathbf{W}^T\mathbf{A}^{-1}\mathbf{V})^{-1}\mathbf{W}^T\mathbf{A}^{-1}$
	†	$-\ \mathbf{A}^{-1}\mathbf{V\,D}(\mathbf{D}+\mathbf{D\,W}^T\mathbf{A}^{-1}\mathbf{V\,D})^{-1}\mathbf{D\,W}^T\mathbf{A}^{-1}$
$\mathbf{U}_I\,\Delta\mathbf{A}_{IJ}\,\mathbf{U}_J^{\,T}$		$-\ \mathbf{A}^{-1}\mathbf{U}_I\,\Delta\mathbf{A}_{IJ}(1+\mathbf{U}_J^{\,T}\mathbf{A}^{-1}\mathbf{U}_I\,\Delta\mathbf{A}_{IJ})^{-1}\mathbf{U}_J^{\,T}\mathbf{A}^{-1}$
(unit vectors selecting proper variable elements)		$-\ \mathbf{A}^{-1}\mathbf{U}_I(1+\Delta\mathbf{A}_{IJ}\,\mathbf{U}_J^{\,T}\mathbf{A}^{-1}\mathbf{U}_I)^{-1}\Delta\mathbf{A}_{IJ}\,\mathbf{U}_J^{\,T}\mathbf{A}^{-1}$
	††	$-\ \mathbf{A}^{-1}\mathbf{U}_I(\Delta\mathbf{A}_{IJ}^{\,-1}+\mathbf{U}_J^{\,T}\mathbf{A}^{-1}\mathbf{U}_I)^{-1}\mathbf{U}_J^{\,T}\mathbf{A}^{-1}$
	††	$-\ \mathbf{A}^{-1}\mathbf{U}_I\Delta\mathbf{A}_{IJ}(\Delta\mathbf{A}_{IJ}+\Delta\mathbf{A}_{IJ}\mathbf{U}_J^{\,T}\mathbf{A}^{-1}\mathbf{U}_I\Delta\mathbf{A}_{IJ})^{-1}\Delta\mathbf{A}_{IJ}\mathbf{U}_J^{\,T}\mathbf{A}^{-1}$
$\mathbf{v\,w}^T\,\Delta\phi$ (rank 1 charge)		$-\ \dfrac{\mathbf{A}^{-1}\mathbf{v\,w}^T\mathbf{A}^{-1}}{\Delta\phi^{-1}+\mathbf{w}^T\mathbf{A}^{-1}\mathbf{v}}$

† where **D** is a nonsingular square matrix.

†† where $\Delta\mathbf{A}_{IJ}$ is a nonsingular square matrix.

TABLE VIII

FORMULAS FOR THE COMPUTATION OF LARGE CHANGES

WHEN $\mathbf{A}^{-1}$ IS INVOLVED AND WHEN $r_1 \geq r_2$

	Identification		Formula	Definition of **S** or **s**
	$\Delta(\mathbf{A}^{-1})$		$-\mathbf{P}_V \mathbf{D} \mathbf{S} \mathbf{Q}_W{}^T$	$\mathbf{H}_1 \mathbf{S} = \mathbf{1}$
	$\Delta(\mathbf{b}^T \mathbf{A}^{-1})$		$-\mathbf{s}^T \mathbf{Q}_W{}^T$	$\mathbf{H}_1{}^T \mathbf{s} = \mathbf{D}^T \mathbf{V}^T \mathbf{q}_b$
	$\Delta(\mathbf{A}^{-1} \mathbf{c})$		$-\mathbf{P}_V \mathbf{D} \mathbf{s}$	$\mathbf{H}_2 \mathbf{s} = \mathbf{W}^T \mathbf{p}_c$
	$\Delta(\mathbf{b}^T \mathbf{A}^{-1} \mathbf{c})$		$-\mathbf{s}^T \mathbf{Q}_W{}^T \mathbf{c}$	$\mathbf{H}_1{}^T \mathbf{s} = \mathbf{D}^T \mathbf{V}^T \mathbf{q}_b$
†	$\Delta(\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})$	(1)	$-\mathbf{S}^T \mathbf{Q}_W{}^T \mathbf{C}$	$\mathbf{H}_1{}^T \mathbf{S} = \mathbf{D}^T \mathbf{V}^T \mathbf{Q}_B$
		(2)	$-\mathbf{B}^T \mathbf{P}_V \mathbf{D} \mathbf{S}$	$\mathbf{H}_2 \mathbf{S} = \mathbf{W}^T \mathbf{P}_C$
		(3)	$-\mathbf{Q}_B{}^T \mathbf{V} \mathbf{D} \mathbf{S}$	$\mathbf{H}_1 \mathbf{S} = \mathbf{Q}_W{}^T \mathbf{C}$
		(4)	$-\mathbf{B}^T \mathbf{P}_V \mathbf{D} \mathbf{S} \mathbf{Q}_W{}^T \mathbf{C}$	$\mathbf{H}_1 \mathbf{S} = \mathbf{1}$
		(5)	$-\mathbf{Q}_B{}^T \mathbf{V} \mathbf{D} \mathbf{S} \mathbf{Q}_W{}^T \mathbf{C}$	$\mathbf{H}_1 \mathbf{S} = \mathbf{1}$

where $\mathbf{H}_1 = (\mathbf{1} + \mathbf{Q}_W{}^T \mathbf{V} \mathbf{D})$, $\mathbf{H}_2 = (\mathbf{1} + \mathbf{W}^T \mathbf{P}_V \mathbf{D})$

† Table X can be used as a guide to select among (1) to (5) by the minimum FBS criterion.

TABLE IX

FORMULAS FOR THE COMPUTATION OF LARGE CHANGES

WHEN $\mathbf{A}^{-1}$ IS INVOLVED AND WHEN $r_1 < r_2$

	Identification		Formula	Definition of S or s
	$\Delta(\mathbf{A}^{-1})$		$-\mathbf{P_V S D Q_W}^T$	$\mathbf{H_2 S = 1}$
	$\Delta(\mathbf{b}^T \mathbf{A}^{-1})$		$-\mathbf{s}^T \mathbf{D Q_W}^T$	$\mathbf{H_1}^T \mathbf{s} = \mathbf{V}^T \mathbf{q_b}$
	$\Delta(\mathbf{A}^{-1} \mathbf{c})$		$-\mathbf{P_V s}$	$\mathbf{H_2 s = D W}^T \mathbf{p_c}$
	$\Delta(\mathbf{b}^T \mathbf{A}^{-1} \mathbf{c})$		$-\mathbf{b}^T \mathbf{P_V s}$	$\mathbf{H_2 s = D W}^T \mathbf{p_c}$
†	$\Delta(\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})$	(1)	$-\mathbf{B}^T \mathbf{P_V S}$	$\mathbf{H_2 S = D W}^T \mathbf{P_C}$
		(2)	$-\mathbf{S}^T \mathbf{D Q_W}^T \mathbf{C}$	$\mathbf{H_1}^T \mathbf{S} = \mathbf{V}^T \mathbf{Q_B}$
		(3)	$-\mathbf{S}^T \mathbf{D W}^T \mathbf{P_C}$	$\mathbf{H_2}^T \mathbf{S} = \mathbf{P_V}^T \mathbf{B}$
		(4)	$-\mathbf{B}^T \mathbf{P_V S D Q_W}^T \mathbf{C}$	$\mathbf{H_2 S = 1}$
		(5)	$-\mathbf{B}^T \mathbf{P_V S D W}^T \mathbf{P_C}$	$\mathbf{H_2 S = 1}$

where $\mathbf{H_1} = (\mathbf{1} + \mathbf{D Q_W}^T \mathbf{V})$, $\mathbf{H_2} = (\mathbf{1} + \mathbf{D W}^T \mathbf{P_V})$

† Table XI can be used as a guide to select among (1) to (5) by the minimum FBS criterion.

TABLE X

MAJOR COMPUTATIONAL EFFORT FOR CALCULATING $\Delta(\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})$

BY FORMULAS IN TABLE VIII WHERE $r_1 \geq r_2$

Category	Corresponding Case in Table VIII	The n x n System Represented By **A**	The r_2 x r_2 System Represented By $\mathbf{H}_1$ or $\mathbf{H}_2$
No. of LU Factorizations	(1) - (5)	1	1
No. of FBS	(1)	$m' + r_2$	m'
	(2)	$n' + r_1$	n'
	(3)	$m' + r_2$	n'
	(4)	$r_1 + r_2$	r_2
	(5)	$m' + r_2$	r_2

TABLE XI

MAJOR COMPUTATIONAL EFFORT FOR CALCULATING $\Delta(\mathbf{B}^T \mathbf{A}^{-1} \mathbf{C})$

BY FORMULAS IN TABLE IX WHERE $r_1 < r_2$

Category	Corresponding Case in Table IX	The n x n System Represented by **A**	The r_1 x r_1 System Represented by $\mathbf{H}_1$ or $\mathbf{H}_2$
No. of LU Factorizations	(1) - (5)	1	1
No. of FBS	(1)	$n' + r_1$	n'
	(2)	$m' + r_2$	m'
	(3)	$n' + r_1$	m'
	(4)	$r_2 + r_1$	r_1
	(5)	$n' + r_1$	r_1

TABLE XII

OPERATIONAL COUNTS FOR $\Delta(\mathbf{A}^{-1})$

n,r_1,r_2	By Direct Method	By Formula (47)		By Formula (64)
		Variables Change for the First Time	Variables Change for the Second Time and On	
10,1,1	1000	323	113	330
10,3,2	1000	852	292	700
10,5,3	1000	1517	567	1110
10,4,4	1000	1712	752	1440
10,6,4	1000	2136	896	1520
10,6,5	1000	2625	1225	1900

TABLE XIII

OPERATIONAL COUNTS FOR $\Delta(\mathbf{A}^{-1}\mathbf{c})$

n, r_1, r_2	By Direct Method	By Formula (47)	
		Variables Change for the First Time	Variables Change for the Second Time and On
10,1,1	430	133	13
10,3,2	430	434	54
10,5,3	430	807	127
10,4,4	430	756	156
10,6,4	430	1096	216
10,6,5	430	1255	305
10,6,6	430	1438	418
10,7,6	430	1650	470
4,3,2	36	116	36
7,2,2	161	172	32

MATRIX [A]										VECTOR [B]
1.0	5.0	5.0	1.0	5.0	2.0	1.0	1.0	7.0	2.0	35.0
2.0	3.0	3.0	7.0	0.0	4.0	3.0	6.0	8.0	3.0	32.0
3.0	0.0	2.0	4.0	2.0	6.0	4.0	4.0	9.0	7.0	16.0
6.0	1.0	2.0	5.0	2.0	3.0	3.0	7.0	3.0	5.0	51.0
8.0	1.0	2.0	2.0	4.0	4.0	6.0	8.0	4.0	8.0	42.0
4.0	1.0	6.0	7.0	3.0	5.0	7.0	3.0	5.0	3.0	19.0
7.0	0.0	6.0	5.0	9.0	4.0	8.0	9.0	2.0	9.0	34.0
2.0	0.0	4.0	2.0	2.0	5.0	3.0	5.0	4.0	3.0	71.0
3.0	2.0	0.0	1.0	5.0	3.0	4.0	2.0	3.0	1.0	36.0
4.0	2.0	4.0	4.0	6.0	2.0	9.0	6.0	1.0	7.0	61.0

SOLUTIONS BEFORE LARGE CHANGE :

VECTOR [X]
-8.89217
39.80097
-3.00067
2.31014
-5.40544
48.42778
-12.11626
-3.61726
-32.93004
16.99799

Fig. 1(a) The original linear system and its solutions. **A** is a 10x10 matrix containing parameters of the system. **b** is the excitation vector. **x** is the solution vector.

MATRIX [V]

0.0	0.0	0.0
1.0	0.0	0.0
0.0	0.0	0.0
0.0	0.0	0.0
0.0	1.0	0.0
0.0	0.0	0.0
0.0	0.0	0.0
0.0	0.0	0.0
0.0	0.0	1.0
0.0	0.0	0.0

MATRIX [W]

0.0	0.0
0.0	0.0
1.0	0.0
0.0	0.0
0.0	0.0
0.0	1.0
0.0	0.0
0.0	0.0
0.0	0.0
0.0	0.0

MATRIX [PV]

-.03684	.19072	-.00936
-.30799	-.26526	.04838
.00454	.09406	-.15865
-.04645	-.23600	.02949
.02608	-.09579	.12002
-.48948	-.43199	.13012
.18919	.22984	.01487
.27060	.13754	.00846
.36321	.32717	-.02238
-.27658	-.20670	-.09789

VECTOR [RHS]

-3.00067

48.42778

Fig. 1(b) Matrices **V**, **W**, $\mathbf{P_V}$ and vector **RHS**, where $\mathbf{P_V}$ is the solution of $\mathbf{A}\ \mathbf{P_V} = \mathbf{V}$ and $\mathbf{RHS} = \mathbf{W}^T \mathbf{x}$.

MATRIX [D]

2.00000	3.00000
4.00000	5.00000
2.00000	3.00000

MATRIX [H]

1.06802	.00797
-2.44666	-2.23801

VECTOR [S]

-2.66983

-18.72004

SOLUTIONS AFTER LARGE CHANGE :

VECTOR [X]

8.15496

-3.82546

-2.66983

-23.34277

-6.40995

-18.72004

24.40133

27.88727

22.14824

-27.58607

Fig. 1(c) Results corresponding to a particular change of variable parameters represented by **D**. **H** repesents $(1 + \mathbf{W}^T \mathbf{A}^{-1} \mathbf{V} \mathbf{D})$ and **s** is the solution of the smaller system $\mathbf{H}\,\mathbf{s} = \mathbf{W}^T \mathbf{x}$.

VARIABLES CHANGE FOR ANOTHER TIME CAUSING THE CHANGE OF [D].
BUT [V] AND [W] ARE THE SAME AS THOSE WITH THE PREVIOUS [D].

MATRIX [D]

6.00000	7.00000
5.00000	4.00000
3.00000	4.00000

MATRIX [H]

1.02160	-.22657
-4.70646	-3.63383

VECTOR [S]

-4.57788
-7.39775

SOLUTIONS AFTER ANOTHER LARGE CHANGE :

VECTOR [X]

-2.20815
3.56798
-4.57738
-12.47901
-3.16642
-7.39775
15.58395
25.41279
12.05534
-20.00992

Fig. 1(d) Results corresponding to another change of variable parameters. H and s are similarly defined to those in Fig. 1(c).

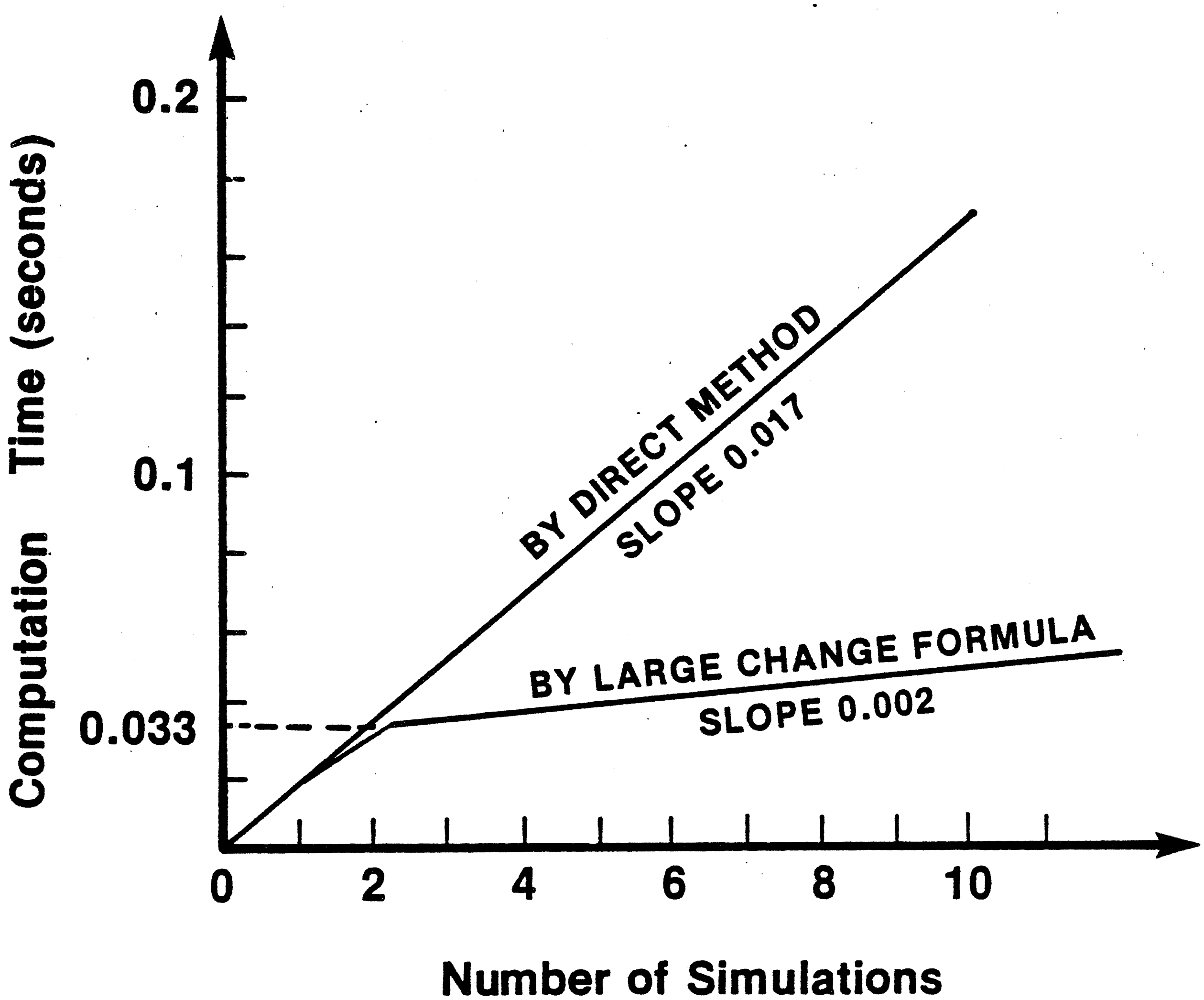

Fig. 2 Computational time for evaluating large change effects of the linear system of the example. (The CDC 170/815 System is used. Compilation time is not included.)