An extreme point result for convexity, concavity and monotonicity of parameterized linear equation solutions

Ashwin Ganesan a, Sheila R. Ross a, B. Ross Barmish b,∗

aEECS Department, University of California-Berkeley, Berkeley, CA, USA
bECE Department, University of Wisconsin-Madison, Madison, WI 53706, USA

Received 13 November 2003; accepted 6 April 2004
Submitted by H. Schneider

Abstract

In many applications, it is useful to know how the solution to a set of simultaneous linear equations depends on parameters θ entering into the coefficients. To this end, this paper addresses the classical equation \( Ax = b \) with \( n \times n \) matrix \( A = A(\theta) \) and \( n \times 1 \) vector \( b = b(\theta) \) depending on an \( m \)-tuple of parameters \( \theta \) with components \( \theta_i \) entering in a rank-one manner. Given such a system, the following problems are considered: For solution component \( x_i(\theta) \) and parameter \( \theta_j \), determine if the first and second order partial derivatives of \( x_i \) with respect to \( \theta_j \) are of one sign for all \( \theta \) in a prescribed hypercube \( \Theta_r \) of radius \( r \geq 0 \); i.e., we determine which components enter the solution either monotonically, convexly or concavely. In this paper, we provide extreme point results for these problems. Namely, we need only check the sign of three specially constructed multilinear functions at the extreme points (vertices) of \( \Theta_r \) in order to ascertain whether the desired one-sign condition is satisfied over the entire hypercube. Central to the proof of extremality is a special “multilinear factorization” of the partial derivatives of \( x_i(\theta) \). This leads to a simple method to compute the so-called radii of convexity, concavity and monotonicity.

© 2004 Elsevier Inc. All rights reserved.

∗ Corresponding author.

E-mail addresses: ganesan@eecs.berkeley.edu (A. Ganesan), sross@eecs.berkeley.edu (S.R. Ross), barmish@engr.wisc.edu (B.R. Barmish).

0024-3795/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
1. Introduction

The focal point of this paper is the dependence of the solution of the linear matrix equation $Ax = b$ on parameters $\theta = (\theta_1, \theta_2, \ldots, \theta_m)$ entering into $A$ and $b$. In many systems science applications, when the solution components $x_i = x_i(\theta)$ are proven to be convex, concave or monotonic, we facilitate the solution of some underlying problem of interest. To illustrate, suppose $x(\theta)$ corresponds to the gain of a system and is a monotonically increasing continuous function of a parameter $\theta \in [\theta_{\min}, \theta_{\max}]$. Then, it follows that the interval of possible gains, call it $\mathcal{X}$, is determined by the two extreme points for $\theta$; i.e., $\mathcal{X} = [x(\theta_{\min}), x(\theta_{\max})]$.

Examples of a similar sort can also be given in a stochastic context. In many cases, when it can be established that $\theta$ enters a performance measure in a convex or concave manner, it becomes possible to carry out a so-called robust Monte Carlo simulation; e.g., see [1,2]. Finally, a third motivating example is obtained when the components of $\theta$ correspond to design parameters and $x_i(\theta)$ is some system quantity to be minimized. Now, convex dependence on $\theta$ facilitates computation of some optimal setting $\theta = \theta^*$ for the system.

To proceed more formally, we consider the set of linear equations $Ax = b$ with $n \times n$ matrix $A \doteq A(\theta)$ and $n \times 1$ vector $b \doteq b(\theta)$ depending on the $m$-tuple of parameters

$\theta \doteq (\theta_1, \theta_2, \ldots, \theta_m)$

which are constrained to lie in a hypercube $\Theta_r$ of radius $r \geq 0$. The objective is to determine if $A(\theta)$ is invertible for all $\theta \in \Theta_r$, and if so, whether the $i$th solution component $x_i(\theta)$ of

$x(\theta) = A^{-1}(\theta)b(\theta)$

is convex, concave, or monotonic with respect to individual components $\theta_j$.

1.1. Rank-one uncertainty structures

The results to follow apply when the pair $(A(\theta), b(\theta))$ has a so-called rank-one uncertainty structure. By this, we mean the following: First, each parameter $\theta_j$ enters into either $A(\theta)$ or $b(\theta)$ but not both. Second, when parameter $\theta_j$ enters into $A(\theta)$, there exist a matrix function $A_0 = A_0(\theta)$ and non-zero constant vectors $d_j, e_j \in \mathbb{R}^n$ such that

$A(\theta) \doteq A_0 + \theta_j d_j e_j^T$
with \( A_0(\theta) \) depending on \( \theta_k \) for \( k \neq j \). Third, when \( \theta_j \) enters into \( b(\theta) \), there exists a vector function \( b_0 = b_0(\theta) \) and a non-zero constant vector \( b_j \in \mathbb{R}^n \) such that

\[
b(\theta) = b_0 + \theta_j b_j
\]

with \( b_0(\theta) \) depending on \( \theta_k \) for \( k \neq j \).

### 1.2. Motivating example

To motivate the results presented in the sequel, it is noted that any planar resistive network has the following property: With resistances \( R_j \) associated with the parameters \( \theta_j \) in this paper, the resulting linear equations, corresponding to application of Kirchhoff’s laws, has a rank-one uncertainty structure. To be more specific, we consider the network in Fig. 1 and make the following identification between circuit variables and linear algebra variables: The \( i \)th solution variable \( x_i \) is taken to be the current \( I_i \) for \( i = 1, 2, 3, 4 \) and \( x_5 = V_{\text{out}} \). Now, a straightforward computation yields

\[
x_5(\theta) = \frac{N(\theta)}{D(\theta)},
\]

where the numerator and denominator, \( N(\theta) \) and \( D(\theta) \), are given by

\[
N(\theta) = -200(10\theta_2 + 10\theta_4 + 10\theta_3 + \theta_3\theta_4) \\
- 350(\theta_1\theta_4 + \theta_2\theta_4 + \theta_2\theta_3 + \theta_1\theta_2 + \theta_1\theta_3) - 30(\theta_2\theta_3\theta_4 + \theta_1\theta_2\theta_4),
\]

\[
D(\theta) = 50(2\theta_1\theta_4 + 10\theta_4 + \theta_2\theta_4 - 10\theta_3 - \theta_2\theta_3) - 50(\theta_2\theta_1 - \theta_1\theta_3).
\]

As seen when this example is revisited in Section 4.1, the derivatives of \( x_5(\theta) \) admit a special factorization in terms of multilinear functions of \( \theta \). As a consequence, it

---

Fig. 1. Example circuit.
becomes easy to classify the \( \theta_j \) into various categories; i.e., it is seen that \( x_5 \) is convex and monotonically decreasing in some \( \theta_j \), concave and monotonically increasing in other \( \theta_j \).

1.3. Monotonicity and convexity considerations

In the sequel, as a preliminary step, we seek to determine whether \( A(\theta) \) is non-singular for all \( \theta \) in the hypercube \( \Theta_r \). When this condition is satisfied, the main objective of this paper is to provide conditions under which the signs of the first and second partial derivatives of the solution variable of interest \( x_i(\theta) \) remain constant, either positive or negative, over \( \Theta_r \). To this end, the center of the hypercube \( \Theta_r \), the so-called nominal parameter vector, is denoted by

\[
\theta^0 = (\theta^0_1, \theta^0_2, \ldots, \theta^0_m),
\]

and we assume

\[
x^0 = A^{-1}(\theta^0)b(\theta^0)
\]

exists and that the partial derivatives

\[
\nabla_{ij}(\theta) = \frac{\partial x_i(\theta)}{\partial \theta_j}, \quad \nabla^2_{ij}(\theta) = \frac{\partial^2 x_i(\theta)}{\partial \theta_j^2}
\]

are non-zero at \( \theta = \theta^0 \). Then, the componentwise monotonicity problem of interest is to determine if the sign of \( \nabla_{ij}(\theta) \) remains invariant over \( \Theta_r \). For the second derivative the componentwise convexity and concavity problems of interest are to determine if the sign of \( \nabla^2_{ij}(\theta) \) remains invariant over \( \Theta_r \). For example, if \( \nabla_{ij}(\theta) < 0 \) and \( \nabla^2_{ij}(\theta) > 0 \) for \( \theta \in \Theta_r \), it follows that \( x_i(\theta) \) is strictly convex and monotonically decreasing in component \( \theta_j \).

1.4. Radii of interest

We first define the radius of non-singularity

\[
r^{\text{NS}} = \sup \{ r : A^{-1}(\theta) \text{ exists for all } \theta \in \Theta_r \}
\]

and note the following: With \( A(\theta) \) having rank-one uncertainty structure, it can readily be shown that \( \det A(\theta) \) is a multilinear affine function of \( \theta \). That is, if all components of \( \theta \) are held fixed except for \( \theta_k \), then \( \det A(\theta) \) is affine linear in \( \theta_k \). Henceforth, for simplicity of exposition, we refer to such multilinear affine functions as being “multilinear.” Moreover, using well-known results on multilinear functions, for example see [3], non-singularity of \( A(\theta) \) over \( \Theta_r \) can be directly determined by checking to see that \( \det A(\theta) \) has the same sign at all extreme points (vertices) of the hypercube \( \Theta_r \).
We can now define three additional radii of interest. First, the radius of strict convexity is given by

$$r_{ij}^C = \sup \{ r \leq r^{NS} : \nabla^2_{ij}(\theta) > 0 \text{ for all } \theta \in \Theta_r \}.$$ 

Similarly, a radius of strict concavity is defined using the condition $\nabla^2_{ij} < 0$ above. Finally, the radius of monotonicity is defined as

$$r_{ij}^M = \sup \{ r \leq r^{NS} : \text{ either } \nabla_{ij}(\theta) > 0 \text{ for all } \theta \in \Theta_r \text{ or } \nabla_{ij}(\theta) < 0 \text{ for all } \theta \in \Theta_r \}.$$ 

1.5. Rank-one implications

Recalling the standing assumption that $(A(\theta), b(\theta))$ has a rank-one uncertainty structure, one of the main results of this paper is that the radii of convexity, concavity and monotonicity are obtained via evaluation of three specially constructed multilinear functions at the extreme points of the hypercube $\Theta_r$. In proving this result, the technical novelty of the paper resides in the way we handle the fact that the multilinear dependence of $x_i$ on $\theta$ is destroyed when partial derivatives are taken in the analysis of convexity and monotonicity. To this end, we prove that these derivatives admit a special factorization over the range of non-singularity. As seen in Section 2, these derivatives can always be expressed as a product of multilinear functions of $\theta$ and a one-sign function of $\theta$.

2. Main results

In this section, the main results of this paper are provided. Namely, we provide a multilinear factorization of the derivatives of the solution variables $x_i(\theta)$ and an extreme point condition which is necessary and sufficient for satisfaction of the desired one-sign derivative conditions over $\Theta_r$.

2.1. Preliminaries

To simplify the exposition to follow, it is first noted that if a parameter $\theta_j$ enters into $b(\theta)$, since $x(\theta) = A^{-1}(\theta)b(\theta)$, the $i$th component $x_i(\theta)$ is an affine linear function $\theta_j$. Hence, writing $b(\theta) = b_0 + \theta_j b_1$ with $b_0$ depending on $\theta_k$ with $k \neq j$, the two partial derivatives of interest are trivially given by

$$\frac{\partial x_i}{\partial \theta_j} = \frac{\eta_i^T b_1}{\det A(\theta)}, \quad \frac{\partial^2 x_i}{\partial \theta_j^2} = 0.$$
where \( \eta_i \) denotes the unit vector in the \( i \)th coordinate direction. Hence, the theorem to follow addresses the more difficult case when \( \theta_j \) is a parameter entering into \( A(\theta) \).

To this end, with \( A(\theta) \) expressed in the rank-one form

\[
A(\theta) = A_0 + \sum_{k=1}^{m} \theta_k d_k e_k^T,
\]

with \( d_k \) and \( e_k \) being fixed non-zero vectors in \( \mathbb{R}^n \), for \( \theta_j \) being the parameter of interest, we define

\[
A_j(\theta) = A_0 + \sum_{k \neq j}^{m} \theta_k d_k e_k^T,
\]

and a triple of multilinear functions given by

\[
\begin{align*}
g_1(\theta) &= (\det A_j) e_j^T A_j^{-1} d_j, \\
g_2(\theta) &= (\det A_j) \eta_i^T A_j^{-1} d_j, \\
g_3(\theta) &= 2(\det A_j) e_j^T A_j^{-1} b.
\end{align*}
\]

In the theorem below, when we refer to a function \( g(\theta) \) as being sign-invariant for \( \theta \in \Theta_r \), we mean that either \( g(\theta) > 0 \) for all \( \theta \in \Theta_r \) or \( g(\theta) < 0 \) for all \( \theta \in \Theta_r \). When \( g(\theta) \) is continuous such a condition is equivalent to \( g(\theta) \) being non-vanishing over \( \Theta_r \).

2.2. Extremality Theorem

Consider the parameterized linear equation \( A(\theta)x = b(\theta) \) with rank-one uncertainty structure and parameter \( \theta_j \) entering into \( A(\theta) \). For solution component \( x_i(\theta) \), assume that the partial derivatives \( \nabla_{ij}(\theta) \) and \( \nabla_{ij}^2(\theta) \) are non-zero for \( \theta = \theta^0 \) and that \( A^{-1}(\theta) \) exists for all \( \theta \in \Theta_r \). Then it follows that the first partial derivative \( \nabla_{ij}(\theta) \) is sign-invariant for \( \theta \in \Theta_r \) if and only if the extreme point evaluations \( g_1(\theta^k) \) and \( g_2(\theta^k) \) are sign invariant. The same result holds for the second partial derivative \( \nabla_{ij}^2(\theta) \) with sign invariance required for \( g_1(\theta^k) \), \( g_2(\theta^k) \) and \( g_3(\theta^k) \). (See Section 3 for proof.)

2.3. Factorization Lemma

Consider the parameterized linear equation \( A(\theta)x = b(\theta) \) with rank-one uncertainty structure and parameter \( \theta_j \) entering into \( A(\theta) \). Then, the first and second partial derivatives of \( x_i(\theta) \) with respect to \( \theta_j \) admit factorizations of the form

\[
\frac{\partial x_i}{\partial \theta_j} = \frac{g_1(\theta) g_2(\theta)}{\det A^2(\theta)}.
\]
\[ \frac{\partial^2 x_i}{\partial \theta_j^2} = \frac{g_1(\theta)g_2(\theta)g_3(\theta)}{\det A_j(\theta)} \]

with \( g_1(\theta), g_2(\theta) \) and \( g_3(\theta) \) being the multilinear functions defined in Section 2.1.

3. Proof of main results

This section is devoted to the proof of Theorem 2.2 and Lemma 2.3. It may be skipped by the reader interested solely in the application of the main results. We first prove the Factorization lemma.

3.1. Proof of Factorization Lemma 2.3

For the multi-parameter system with rank-one uncertainty structure and notation as given in Section 2.1, noting that

\[ A(\theta) = A_j(\theta) + \theta_j d_j e_j^T, \]

to obtain the solution \( x_i(\theta) \), we first invert \( A(\theta) \) using the Sherman–Morrison equation \([4]\). Temporarily suppressing the dependence on \( \theta \) in the notation, we obtain

\[ A^{-1} = A_j^{-1} - \theta_j \frac{(A_j^{-1} d_j)(e_j^T A_j^{-1})}{1 + \theta_j e_j^T A_j^{-1} d_j}. \]

Now, recalling that \( \eta_i \) is a unit vector in the \( i \)th coordinate direction and noting that the corresponding solution component is \( x_i = \eta_i^T A^{-1} b \), we differentiate with respect to \( \theta_j \) and obtain

\[ \frac{\partial x_i}{\partial \theta_j} = \frac{(\eta_i^T A_j^{-1} d_j)(e_j^T A_j^{-1} b)}{(1 + \theta_j e_j^T A_j^{-1} d_j)^2}. \]

Differentiating a second time yields

\[ \frac{\partial^2 x_i}{\partial \theta_j^2} = \frac{2(e_j^T A_j^{-1} d_j)(\eta_i^T A_j^{-1} d_j)(e_j^T A_j^{-1} b)}{(1 + \theta_j e_j^T A_j^{-1} d_j)^3}. \]

Rewriting both partial derivatives in terms of \( g_1, g_2 \) and \( g_3 \), we obtain

\[ \frac{\partial x_i}{\partial \theta_j} = \frac{g_1(\theta)g_2(\theta)}{(\det A_j)^2(1 + \theta_j e_j^T A_j^{-1} d_j)^2}, \]

\[ \frac{\partial^2 x_i}{\partial \theta_j^2} = \frac{g_1(\theta)g_2(\theta)g_3(\theta)}{(\det A_j)^3(1 + \theta_j e_j^T A_j^{-1} d_j)^3}. \]
Using the identities
\[
\det A(\theta) = \det \left[ A_j + \theta_j d_j e_j^T \right]
= (\det A_j) \left( 1 + \theta_j e_j^T A_j^{-1} d_j \right),
\]
we find
\[
\frac{\partial x_i}{\partial \theta_j} = \frac{g_1(\theta) g_2(\theta)}{\det A^2(\theta)},
\]
\[
\frac{\partial^2 x_i}{\partial \theta_j^2} = \frac{g_1(\theta) g_2(\theta) g_3(\theta)}{\det A^3(\theta)}.
\]

Finally, to establish that each of the functions \(g_i(\theta)\) is multilinear, we note the following: Since each parameter enters \(A(\theta)\) in a rank-one manner, if we hold all components of \(\theta\) fixed except for \(\theta_k\), the cofactors of \(A_j\) are affine linear functions of \(\theta_k\). Keeping in mind that \(b, \eta_i, d_j\) and \(e_j\) are all independent of the particular \(\theta_k\) that enter into \(A(\theta)\), it is clear that the functions \(g_i(\theta)\) are multilinear in \(\theta\).

3.2. Proof of Extremality Theorem 2.2

The proof of this theorem follows almost immediately from the Factorization Lemma. Indeed, noting that \(r \leq r_{NS}\), since \(\det A(\theta)\) and \(\det A^2(\theta)\) are continuous functions which do not vanish over \(\Theta_r\), each of these determinants has one sign. Assuming, without loss of generality, that these determinants are positive, it follows that the sign of the derivatives of \(x_i(\theta)\) are determined by the sign of the multilinear functions \(g_i(\theta)\). The proof is now completed by invoking the well-known fact that a multilinear function \(g_i(\theta)\) on a hypercube is positive if and only if the extreme point evaluations \(g_i(\theta^k)\) are positive; e.g., see [5].

4. Illustrative examples

In this section, we provide two examples illustrating the application of Theorem 2.2. The circuit example from Section 1.2 is revisited and the analysis is completed via application of the theorems. Then, a second example is considered involving an interval matrix \(A(\theta)\). This example demonstrates the efficacy of the theorem even when \(\theta\) has high dimension; i.e., for the \(3 \times 3\) case considered, \(\theta\) is nine-dimensional. Note that it is trivial to verify that every interval matrix \(A(\theta)\) satisfies the rank-one structural requirement of this paper.

4.1. Example (resistive circuit)

The application of Theorem 2.2 is now demonstrated for the resistive network in Section 1.2. With the identification between linear algebra variables and circuit
variables already indicated, the output voltage $x_5$ is obtained via solution of the linear mesh equations

$$
\begin{bmatrix}
30 & -10 & -10 & -10 & 0 \\
-10 & 10 + \theta_2 + \theta_1 & -\theta_1 & 0 & 0 \\
-10 & -\theta_1 & 15 + \theta_3 + \theta_4 & -5 & 0 \\
0 & 0 & -5 & 15 + \theta_4 & 0 \\
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5 \\
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
0 \\
0 \\
0 \\
0 \\
\end{bmatrix}
$$

Assuming resistances $\theta_j$ with nominal values $\theta_1^0 = 10 \, \Omega$, $\theta_2^0 = 6 \, \Omega$, $\theta_3^0 = 10 \, \Omega$, and $\theta_4^0 = 10 \, \Omega$, and that each $\theta_j$ lies within $\pm 2 \, \Omega$ of its nominal value, we check for convexity, concavity and monotonicity over the box $\Theta_r = [8, 12] \times [4, 8] \times [8, 12] \times [8, 12]$.

To study the dependence of $x_5$ on each $\theta_j$, each multilinear factor associated with the Factorization Lemma is computed. For $\theta_1$, we obtain

$$
g_1(\theta) = -150\theta_3 - 10\theta_3\theta_4 - 50\theta_4 + 200\theta_2 + 10\theta_2\theta_4,
$$

$$
g_2(\theta) = -500\theta_4 - 500\theta_2 - 150\theta_2\theta_4 - 500\theta_3,
$$

$$
g_3(\theta) = -700(\theta_2 + \theta_3 + \theta_4) - 60(\theta_2\theta_3 + \theta_3\theta_4),
$$

$$
det A(\theta) = 200(10\theta_2 + 10\theta_4 + 10\theta_3 + 3\theta_3\theta_4)
+ 350(\theta_1\theta_4 + \theta_2\theta_4 + \theta_2\theta_3 + \theta_1\theta_2 + \theta_1\theta_3) + 30(\theta_2\theta_3\theta_4 + \theta_1\theta_2\theta_4).
$$

Keeping in mind that the resistor values are positive, it follows by inspection that $\det A(\theta)$ is positive over $\Theta_r$ and that $g_2(\theta)$ and $g_3(\theta)$ are negative over $\Theta_r$. Hence, the problems of convexity, concavity and monotonicity reduce to checking the sign of $g_1(\theta)$ over $\Theta_r$. Now, according to Theorem 2.2, we need only check the sign of $g_1(\theta)$ at each extreme point $\theta^E$ of the hypercube $\Theta_r$. Since $g_1(\theta^E)$ is readily verified to be negative for all such extremes, we conclude that $x_5(\theta)$ is concave and monotonically increasing with respect to $\theta_1$.

Proceeding to study dependence of $x_5$ on the remaining $\theta_j$, the same procedure leads to the conclusion that $x_5(\theta)$ is convex and monotonically decreasing with respect to $\theta_2$ and $\theta_3$ and concave and monotonically increasing with respect to $\theta_4$. It is interesting to note that circuit analysis indicates that the sign of $g_1(\theta)$ also determines the direction of the flow of current through the resistor $R_1$; current flows to the left if this quantity is positive and right if negative. In the context of this example, this lends a physical interpretation to the sign constancy criterion; that is, convexity, concavity or monotonicity of $x_5$ with respect to $\theta_1$ is determined by the direction of current flow through $R_1$.

### 4.2. Interval matrix

As a second illustration of Theorem 2.2, we consider the case when $A(\theta)$ is an interval matrix and seek radii of convexity and concavity. For the system of linear equations
Table 1

<table>
<thead>
<tr>
<th>Component</th>
<th>Dependence of $x_2(\theta)$</th>
<th>$r^{C} _{2j}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>Convex</td>
<td>1.77</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>Convex</td>
<td>2.79</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>Convex</td>
<td>0.64</td>
</tr>
<tr>
<td>$\theta_4$</td>
<td>Concave</td>
<td>1.72</td>
</tr>
<tr>
<td>$\theta_5$</td>
<td>Concave</td>
<td>1.72</td>
</tr>
<tr>
<td>$\theta_6$</td>
<td>Concave</td>
<td>0.64</td>
</tr>
<tr>
<td>$\theta_7$</td>
<td>Concave</td>
<td>1.77</td>
</tr>
<tr>
<td>$\theta_8$</td>
<td>Concave</td>
<td>2.79</td>
</tr>
<tr>
<td>$\theta_9$</td>
<td>Concave</td>
<td>0.64</td>
</tr>
</tbody>
</table>

\[
\begin{pmatrix}
16 + \theta_1 & -18 + \theta_2 & -12 + \theta_3 \\
-11 + \theta_4 & -17 + \theta_5 & 14 + \theta_6 \\
-10 + \theta_7 & 6 + \theta_8 & -13 + \theta_9
\end{pmatrix}\begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix} = \begin{pmatrix}
-7 \\
10 \\
-1
\end{pmatrix},
\]

the radius of non-singularity $r^{NS}$ is determined first. To this end, we first note that $A(0)$ is non-singular. Now, since the rank-one uncertainty structure guarantees that $\det A(\theta)$ is a multilinear function, we may determine $r^{NS}$ by gradually increasing the radius $r$ until the point that $\det A(\theta)$ reaches zero on one of the extreme points of $\Theta_j$. Using this method, we find that $r^{NS} \approx 2.79$.

Now focusing on solution component $x_2(\theta)$, the radii of convexity $r^{C} _{2j}$ are determined using extreme point evaluations as prescribed by Theorem 2.2. As in the case of the radius of non-singularity, the radius $r$ is gradually increased until failure of the convexity test occurs at an extreme point of $\Theta_j$. Via a straightforward computation, we generate the required radii as summarized in Table 1.

5. Some applications of convexity, concavity and monotonicity

Further to Theorem 2.2, there are often important ramifications resulting from convexity, concavity and monotonicity properties. To illustrate, for the case of the resistive network in Section 4.1, the monotonicity results lead to the conclusion that the maximum possible output voltage is obtained by setting $R_1$ and $R_3$ to their upper limits and $R_2$ and $R_4$ to their lower limits. By a similar argument, the minimum possible output voltage is obtained by setting $R_1$ and $R_3$ to their lower limits and $R_2$ and $R_4$ to their upper limits. In summary, two distinguished vertices of the resistor hypercube determine the “envelope” for the output voltage.

5.1. Distributional robustness

A second interesting application of Theorem 2.2 is found in the newly emerging theory of distributionally robust Monte Carlo simulation; e.g., see [2] for a tutorial
survey. Whereas the use of traditional Monte Carlo simulation requires known probability distributions for the uncertain parameters, for example, see [6], distributionally robust Monte Carlo simulation does not.

To provide one illustrative example of a distributional robustness formulation and its connection to the results of this paper, we consider the case when \( \theta \) is a random vector supported in the set \( \Theta_r \). In the theory of distributional robustness, a known mean \( \theta_0 \) for \( \theta \) is assumed and each component \( \theta_j \) are taken to be independent with probability density function \( f_j \) which is unknown except for the fact that it is symmetric and non-increasing with respect to \(|\theta_j - \theta_0^j|\). This defines a class of admissible distributions for \( \theta_j \) which includes the Dirac delta function at \( \theta_j = \theta_0^j \) as a limiting case. Letting \( \mathcal{F} \) denote the resulting class of admissible probability density functions for \( \theta \), for solution component \( x_i(\theta) \), one might seek to find the distributionally robust expected value

\[
\delta^* = \max_{f \in \mathcal{F}} \mathbb{E}[x_i(\theta^f)],
\]

where \( \theta^f \) is the random vector with probability density function \( f \in \mathcal{F} \).

Now, to see how the results in this paper connect with the theory above, it is noted that a maximizing distribution \( f^* \in \mathcal{F} \) leading to \( \delta^* \) is attained as follows:

If \( x_i(\theta) \) is convex in \( \theta_j \), then the distributionally robust expected value is attained with \( f^*_j \) being the uniform distribution over the support interval \([-r, r]\). Similarly, if \( x_i(\theta) \) is concave with respect to \( \theta_j \), then the distributionally robust expected value is attained with \( f^*_j \) being the Dirac delta function.; see [2,7]. To make these ideas more concrete, we revisit the circuit of Section 1.2 for a second time.

![Fig. 2. Convergence of expected value estimates.](image)
5.2. Second revisit of circuit example

To obtain the distributionally robust expected value for the output voltage $x_5(\theta)$, we recall from Section 4.1 that $x_5(\theta)$ is convex in $\theta_2$ and $\theta_3$. Hence, $\mathcal{E}[x_5(\theta)]$ is maximized when these variables are uniformly distributed over their allowed intervals. Similarly, concavity of $x_5(\theta)$ in $\theta_4$ and $\theta_1$ implies that $\mathcal{E}[x_5(\theta)]$ is maximized when these variables have Dirac delta function distribution centered at their nominal values. These maximizing distributions can now be used to estimate $\mathcal{E}^*$ via Monte Carlo sampling. Using over 20,000 samples, Fig. 2 depicts the estimate of $\mathcal{E}^*$ as the number of samples increases. The minimum expected value, obtained with the opposite choices for distributions, is also shown. The value of $\mathcal{E}[x_5(\theta)]$ with each $\theta_j$ uniformly distributed is shown as well; this value underestimates $\mathcal{E}^*$ by approximately 7.5%.

6. Concluding remarks

Central to the main results of this paper is the factorization of the derivatives of the solution of $A(\theta)x = b(\theta)$ into a product of multilinear functions. Although a “user” of this result need not actually perform this factorization, it is fundamental to the researcher trying to extend the results of this paper. By way of future research, it would be of interest to consider the case when the matrix $A(\theta)$ is no longer square and $x(\theta)$ now corresponds to a least squares solution. For example, with $A(\theta)$ being $m \times n$ with rank $m$ for all admissible $\theta$, with least squares solution

$$x^{LS}(\theta) = A^T(\theta)[A(\theta)A^T(\theta)]^{-1}b(\theta),$$

we can again define partial derivatives

$$\nabla_{ij}(\theta) = \frac{\partial x^{LS}}{\partial \theta_j},$$

$$\nabla^2_{ij}(\theta) = \frac{\partial^2 x^{LS}}{\partial \theta_j^2}$$

and consider the extent to which these functions are monotonic, convex or concave. Results along these lines would be important when studying the dependence of the least squares solution on either design variables or uncertain parameters.

Finally, we recall that it was assumed at the outset that each component $\theta_j$ of $\theta$ enters into the linear equations in a rank-one manner. Since the techniques used in this paper do not appear to be generalizable with $\theta_j$ entering in a more complicated way, some entirely new line of attack seems to be required for this more general problem class.
Acknowledgments

Funding for this research was provided by the National Science Foundation under Grant Number ECS-9811051. Preliminary versions of this work are found in conference papers [8,9].

References