4 Robustness to Coefficient Uncertainty

This lecture examines robustness of linear equation solving with respect to perturbations of coefficients of such equations. It develops a matrix version of the small gain theorem, which relates robustness to induced gain bounds, and then generalizes the approach using the notion of a quadratic constraint.

Let $L$ be an $m$-by-$m$ real matrix. Finding the right inverse of $L$ can be interpreted as designing a linear transformation $F$ with input vector $v \in \mathbb{R}^m$ and output $u = Fv \in \mathbb{R}^m$ such that $Lu = v$ for all $v$. So far, the tools introduced in these lectures allow one to study robustness of such linear equation solving with respect to additive perturbations. For example, the L2 (Euclidean norm) gain $\|F\|$ of $F$ quantifies worst-case sensitivity of $u$ with respect to perturbations of $v$, while the Frobenius norm $\|F\|_F$ of $F$ quantifies average sensitivity.

This lecture is concerned with a setup in which the square matrix $L$ is not perfectly known, i.e. it is assumed that $L = L_0 + \Delta$ where $L_0$ is a known nominal value of $L$, and $\Delta$ is the model uncertainty, ranging over a given set $\Delta \in \Delta$ (see Figure 4.1). The questions of interest include establishing right invertibility of $L = L_0 + \Delta$ for all $\Delta \in \Delta$ and quantifying sensitivity of the solution $u$ of $Lu = v$ with respect to $\Delta$.

![Figure 4.1: Uncertainty in the Coefficients](image)

4.1 Small Gain Theorem for Matrices

Consider first the setup on Figure 4.1 in which the set $\Delta$ of all possible $\Delta$ is defined by the condition $\|\Delta\| \leq \delta$, where $\|\Delta\|$ denotes the L2 gain (or "operator norm", or "Euclidean norm induced gain" of $\Delta$), i.e. $\Delta$ is allowed to be any real $m$-by-$m$ matrix with L2 gain not exceeding a given number $\delta > 0$. In this case necessary and sufficient conditions of robust invertibility of $L_0 + \Delta$ are given by the following matrix algebra version of the small gain theorem.
Theorem 4.1 Let $L_0, F_0$ be real $m$-by-$m$ matrices such that $L_0 F_0 = I_m$. Then for every real $\delta > 0$ the following conditions are equivalent:

(a) matrix $L = L_0 + \Delta$ is invertible for every real $m$-by-$m$ matrix $\Delta$ such that $\|\Delta\| \leq \delta$;
(b) $\delta \|F_0\| < 1$.

Moreover, if conditions (a),(b) are satisfied then

$$\| (L_0 + \Delta)^{-1} \| \leq \frac{\|F_0\|}{1 - \delta \|F_0\|}$$

whenever $L^2$ gain of $\Delta$ is not larger than $\delta$.

Proof. To prove that (b) implies (a), note that, as long as $I_m + \Delta F_0$ is invertible, $F_0 = F_0 (I_m + \Delta F_0)^{-1}$ is a right inverse of $L_0 + \Delta$, as

$$(L_0 + \Delta) F_0 = L_0 F_0 + \Delta F_0 = I_m + \Delta F_0.$$  

On the other hand, since $\delta \|F_0\| < 1$ means $|v| > \delta |F_0 v|$ for all $v \neq 0$, we have $|\Delta F_0 v| \leq \delta |F_0 v| < |v|$ and hence $v \neq \Delta F_0 v$ for $v \neq 0$, which means that $I_m + \Delta F_0$ is invertible, as a square matrix with zero null-space.

To prove that (a) implies (b), note that for every pair of vectors $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$, and real number $\delta \geq 0$ such that $\delta |u| \geq |v|$ and $u \neq 0$ there exists an $n$-by-$m$ matrix $\Delta$ such that $L^2$ gain of $\Delta$ is not larger than $\delta$, and $\Delta u = v$. Indeed, one can define $\Delta$ as an appropriately scaled orthogonal projection onto the linear subspace spanned by $v$:

$$\Delta = u (v' v)^{-1} v'$$ has $L^2$ gain $\|D\| = |v|(v' v)^{-1}|u| = |v|/|u|.$$

If $\delta \|F_0\| \geq 1$ then there exists $v \neq 0$ such that $\delta |u| \geq |v|$ for $u = -F_0 v$. Then, for $\Delta$ defined above,

$$(L_0 + \Delta) u = L_0 u + v = L_0 u - L_0 u = 0,$$

which implies that $L_0 + \Delta$ is not invertible.

Finally, assume that conditions (a),(b) are satisfied. For every $u, e$ such that $(L_0 + \Delta) u = e$ let $v, w$ be defined by $u = F_0 v$ and $w = \Delta u$. Then

$$|e| = |v + w| \geq |v| - |w| \geq |v| - \delta |F_0 v| \geq |v| - \delta \|F_0\| \cdot |v| \geq (1 - \delta \|F_0\|)|v|,$$

hence

$$|u| = \|F_0 v\| \leq \|F_0\| \cdot |v| \leq \frac{\|F_0\| \cdot |e|}{1 - \delta \|F_0\|},$$

which proves the upper bound for $\|F\|$. ■
Figure 4.2: Small Gain Theorem: an Interpretation

A useful interpretation of Theorem 4.1 is given on Figure 4.2. On one hand, it points out that solving equation \((L_0 + \Delta)u = h\) can be reduced to solving equation \(u = F_0(h + \Delta u)\). Most importantly, however, it shows that the "well posedness" of the algebraic loop with \(F_0\) and \(\Delta\) in the left side of Figure 4.2 can be guaranteed by making sure that the total "loop gain" \(\|F_0\| \cdot \|\Delta\|\) is strictly less than one.

It is useful to notice that replacing the inequality \(\|\Delta\| \leq \delta\) by the formally more restrictive constraint \(\|\Delta\|_F \leq \delta\) does not allow one to relax the robust invertibility condition \(\|F_0\| < \delta^{-1}\) in Theorem 4.1, (b). This is due to the fact that the graphs

\[
\Gamma_\Delta = \{(u, \Delta u) : \ u \in \mathbb{R}^m, \ \Delta \in \Delta\}
\]

of the uncertain map \(\Delta \in \Delta\) are the same whether \(\Delta\) is defined by bounding the gain or the Frobenius norm.

4.2 Small Gain Theorem for Non-Linear Functions

What happens to the robustness of solving equation \(L_0u + \Delta(u) = v\) when \(\Delta\) is not a linear function?

A simple example shows that condition \(\|\Delta\| \cdot \|F_0\| < 1\) does not guarantee existence of a solution \(u\) in \(L_0u + \Delta(u) = v\) for every \(v \in \mathbb{R}^m\): for

\[
m = 1, \ L_0u = 2u, \ \Delta(u) = \begin{cases} 
0, & u < 1, \\
1, & u \geq 1,
\end{cases}
\]

we have \(F_0 = 0.5\) and \(\|\Delta\| = 1\), hence \(\|\Delta\| \cdot \|F_0\| = 0.5 < 1\), but the equation \(L_0u + \Delta(u) = v\) does not have a solution for \(2 \leq v < 3\) due to the discontinuity of \(\Delta\).

Analysis of the proof of Theorem 4.1 shows that, as long as \(\|\Delta\| \cdot \|F_0\| < 1\), the inequality \(|u| \leq \gamma|v|\) with \(\gamma = \|F_0\|/(1-\|F_0\| \cdot \|\Delta\|)\) is satisfied whenever \(L_0u + \Delta(u) = v\), i.e. a sensitivity bound holds even if well-posedness is not guaranteed.

One way to fix the well-posedness question is to assume that \(\Delta : \ \mathbb{R}^m \mapsto \mathbb{R}^m\) is continuous, in which case the small gain condition \(\|\Delta\| \cdot \|F_0\| < 1\) guarantees existence (though not uniqueness) of a solution \(u\) of \(L_0u + \Delta(u) = v\) for every \(v\), due to the classical Brower’s fixed point theorem, stating that a continuous map \(f\) of a closed disk in \(\mathbb{R}^m\) into itself always always has a ”fixed point” \(f(x) = x\).
It is reasonable to argue that, in the case when $\Delta$ is not linear, the condition $|u| \leq \gamma|v|$ does not really bound the sensitivity of solution $u$ of $L_0u + \Delta(u) = v$ with respect to $v$, as it does not preclude $|u_1 - u_2|$ from being large when $L_0u_1 + \Delta(u_1) = v_1$, $L_0u_2 + \Delta(u_2) = v_2$, and $|v_1 - v_2|$ is small. One fix for this is to switch to the analysis using incremental gains.

In particular, the incremental L2 gain $\|D\|_i$ of a function $\Delta : \mathbb{R}^m \rightarrow \mathbb{R}^n$ can be defined as the minimal upper bound of the ratio $|\Delta(u_1) - \Delta(u_2)|/|u_1 - u_2|$ over all pairs $u_1 \neq u_2$ (naturally, for linear functions, incremental L2 gain equals the standard L2 gain). The proof of Theorem 4.1 can be re-written to show that condition $\|\Delta\|_i \cdot \|F_0\| < 1$ guarantees existence and uniqueness of a solution $u$ of $L_0u + \Delta(u) = v$ for every $v$, and the bound $\|F\|_i \leq \|F_0\|/(1 - \|F_0\| \cdot \|\Delta\|_i)$ for the incremental L2 gain of the resulting inverse function $F$ such that $L_0F(v) + \Delta(F(v)) = v$ for all $v$.

### 4.3 Robustness Analysis with Quadratic Constraints

The key element of robustness analysis performed in the proof of Theorem 4.1 is the idea of replacing the relation $w = \Delta u$, where $\|\Delta\| \leq \delta$ by the inequality $\delta|u| \geq |w|$, after which the proof reduces to manipulations with inequalities. This approach turns out to be very helpful when working with more complex descriptions of the uncertainty set $\Delta = \{\Delta\}$.

Consider, for example, the case when only the upper left $a$-by-$b$ and lower right $b$-by-$a$ blocks of the $m$-by-$m$ matrix $L$ are uncertain ($a + b = m$), so that $L$ has the form $L = L_0 + \Delta$, where $L_0$ is a known matrix,

$$\Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix},$$

matrices $\Delta_1$ and $\Delta_2$ have dimensions $a$-by-$b$ and $b$-by-$a$ respectively, satisfy known gain bounds $\|\Delta_i\| \leq \delta_i$, and are arbitrary otherwise. Then $L = L_0 + \Delta$ can be represented as

$$L = L_0 + B_1\Delta_1C_1 + B_2\Delta_2C_2, \quad \|\Delta_i\| \leq \delta_i,$$

where

$$B_1 = \begin{bmatrix} I_a \\ 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 \\ I_b \end{bmatrix}, \quad C_1 = \begin{bmatrix} I_b & 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0 & I_a \end{bmatrix}.$$  

Accordingly, equation $Lu = h$ can be re-written in the form

$$L_0u + B_1w_1 + B_2w_2 = h, \quad |w_i| \leq \delta_i|C_iu|.$$  

Now the task of proving robust invertibility of $L$ and bounding the gain $\|F\|$ of the inverse $F = L^{-1}$ can be reduced to proving that conditions (4.1) imply inequality $\gamma|h| \geq |u|$ for $\gamma$ as small as possible.

An efficient way of accomplishing this is based on working with quadratic inequalities. For the example under consideration, this means writing the inequalities in (4.1) and the
“inequality-to-be derived” \( \gamma |h| \geq |u| \) in the equivalent form of quadratic constraints

\[
\sigma_i(u, w_1, w_2) = \delta_i^2 |C_i u|^2 - |w_i|^2 \geq 0 \quad (i = 1, 2), \\
\sigma_0(u, w_1, w_2) = \gamma^2 |L_0 u + B_1 w_1 + B_2 w_2|^2 - |u|^2 \geq 0,
\]

and seeking to find a set of Lagrange multipliers \( \tau_i \geq 0 \) such that the inequality

\[
\sigma_0(u, w_1, w_2) \geq \tau_1 \sigma_1(u, w_1, w_2) + \tau_2 \sigma_2(u, w_1, w_2) \quad (4.2)
\]

is satisfied for all \( u \in \mathbb{R}^m, w_1 \in \mathbb{R}^a \) and \( w_2 \in \mathbb{R}^b \).