ECE 598: Interplay between Control and Machine Learning Spring 2019

Lecture 7

Supply Rate Constructions and Quadratic Constraints, Part I

Lecturer: Bin Hu, Date:02/07/2019

As we have already seen in the past lectures, one way to analyze the troublesome element Δ is to replace it with some quadratic supply rate conditions that can be directly plugged in some LMIs for stability analysis.

How to construct supply rate conditions becomes a key issue for such analysis. Fortunately, many such supply rate conditions have already been documented in the controls literature. In today's lecture, we will look at a few basic ones including

1. Small gain,
2. Passivity,
3. Sector bound.

These conditions are also called "quadratic constraints". For simplicity, we will first talk about the pointwise versions of these conditions. Then we will briefly discuss the "integral" versions of these conditions which are the so-called integral quadratic constraints (IQCs).

7.1 Pointwise Quadratic Constraints

Consider a perturbation operator Δ that maps v to w in a static manner, i.e. w_k is completely determined by v_k. The pointwise quadratic constraint just enforces the following inequality on the input/output pair of Δ:

$$\begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix}^\mathsf{T} M \begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix} \leq 0, \tag{7.1}$$

where M is a symmetric matrix, and (w^*, v^*) are typically determined by the fixed points of the feedback interconnection $F_u(G, \Delta)$. The terminology "pointwise" just means that we require the above inequality to hold for all k. Clearly, many supply rate conditions that we have used so far are in the form of such pointwise quadratic constraints. Suppose $v_k - v^* = C(\xi_k - \xi^*)$. Then the quadratic constraint (7.1) just gives the following supply rate condition

$$\begin{bmatrix} \xi_k - \xi^* \\ w_k - w^* \end{bmatrix}^\mathsf{T} \left(\begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix}^\mathsf{T} M \begin{bmatrix} C & 0 \\ 0 & I \end{bmatrix} \right) \begin{bmatrix} \xi_k - \xi^* \\ w_k - w^* \end{bmatrix} \leq 0.$$

For now, we just focus on how to obtain the quadratic constraint (7.1).

7.1.1 Small gain bound

Suppose Δ is bounded in the sense that we have $\|w_k - w^*\| \leq L\|v_k - v^*\|$. The parameter L can be viewed as the input-output gain of the operator Δ. The small gain bound $\|w_k - w^*\| \leq L\|v_k - v^*\|$ is equivalent to the quadratic inequality $\|w_k - w^*\|^2 - L^2\|v_k - v^*\|^2 \leq 0$ which can be rewritten as the following quadratic constraint:

$$\begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix}^\mathsf{T} \begin{bmatrix} -L^2 I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix} \leq 0. \tag{7.2}$$

This is the most commonly-used quadratic constraint. Now let's see a few examples.

- Uncertainty in a multiplicative form: Let Δ map v to w as $w_k = \delta_k v_k$ where δ_k is a matrix changing with k. If we know the Frobenius norm of δ_k is bounded above by L for all k, then we have the small gain bound (7.2) for $(v^*, w^*) = (0, 0)$.
- Gradients of L-smooth functions: Let Δ map v to w as $w_k = \nabla f(v_k)$ where f is L-smooth. Then we have the small gain bound (7.2) holds for any reference point (v^*, w^*) satisfying $w^* = \nabla f(v^*)$.

7.1.2 Passivity

In its simplest form, passivity can be used to describe a function that is in the first and third quadrants. For illustrative purposes, consider a scaler case. Suppose $w_k = \phi(v_k)$ where the function $\phi : \mathbb{R} \to \mathbb{R}$ satisfies $\phi(0) = 0$ and is in the first and third quadrants. Clearly v_k and w_k are both scalers in this case. If $v_k \geq 0$, we have $w_k \geq 0$. If $v_k \leq 0$, we have $w_k \leq 0$. Hence we always have $w_k^\mathsf{T} v_k \geq 0$. This is the basic form of passivity.

A slightly more general form of passivity gives the constraint $(w_k - w^*)^\mathsf{T}(v_k - v^*) \geq 0$ when (v_k, w_k) are vectors and (potentially non-zero) reference points (v^*, w^*) are used. The passivity condition can be rewritten as the following quadratic constraint (verify it!):

$$\begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix}^\mathsf{T} \begin{bmatrix} 0 & -I \\ -I & 0 \end{bmatrix} \begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix} \leq 0. \tag{7.3}$$

Example: Gradients of Convex Functions. Let Δ map v to w as $w_k = \nabla f(v_k)$ where f is a convex function. By definitions, the following inequalities hold for any (v_k, v^*):

$$f(v_k) - f(v^*) \geq \nabla f(v^*)^\mathsf{T}(v_k - v^*)$$
$$f(v^*) - f(v_k) \geq \nabla f(v_k)^\mathsf{T}(v^* - v_k)$$

Summing the above two inequalities directly leads to the passivity condition $(w_k - w^*)^\mathsf{T}(v_k - v^*) \geq 0$. Therefore, gradients of convex functions satisfy the passivity condition.

7.1.3 Sector Bound

Originally sector bound was used to describe a function that is in a sector formed by two lines whose slopes are m and L. First we consider a scaler case. Suppose $w_k = \phi(v_k)$ where the function $\phi : \mathbb{R} \to \mathbb{R}$ satisfies $\phi(0) = 0$ and is in a sector formed by two lines whose slopes are m and L. For simplicity, we assume $L \geq m$. Clearly the sector assumption just ensures $(Lv_k - w_k)^\mathsf{T}(w_k - mv_k) \geq 0$. This is the basic form of the sector bound condition.

Now we can introduce the more general form of the sector bound condition that gives the constraint $\left(L(v_k - v^*) - (w_k - w^*)\right)^\mathsf{T} \left(w_k - w^* - m(v_k - v^*)\right) \geq 0$ when (v_k, w_k) are vectors and the reference points (v^*, w^*) are allowed to be non-zero. The sector bound condition can be rewritten as the following quadratic constraint (verify it!):

$$\begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix}^\mathsf{T} \begin{bmatrix} 2mLI & -(m+L)I \\ -(m+L)I & 2I \end{bmatrix} \begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix} \leq 0. \tag{7.4}$$

Example: Gradients of L-Smooth m-Strongly Convex Functions. Let Δ map v to w as $w_k = \nabla f(v_k)$ where f is L-smooth and m-strongly convex. Then Δ satisfies the sector bound condition (7.4). We have used this condition to prove the linear convergence rate of the gradient method in the previous lectures.

7.2 Integral Quadratic Constraints

In general, Δ is just an operator that maps a sequence $\{v_k\}$ to another sequence $\{w_k\}$. In controls literature, we typically confine Δ to be a causal operator in the sense that w_k is completely determined by $\{v_0, v_1, \ldots, v_k\}$. Here Δ is not static anymore. There may be dynamics involved in Δ. Examples include norm-bounded LTI uncertainty and time-varying delays. For such type of Δ, the pointwise quadratic constraints no longer hold. However, the quadratic constraints may hold when we sum them. Specifically, the integral quadratic constraints (IQCs) just enforce the following inequality for any N,

$$\sum_{k=0}^{N} \begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix}^\mathsf{T} M \begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix} \leq 0, \tag{7.5}$$

Originally the above type of quadratic constraints were developed in continuous-time domain where the quadratic forms are integrated over the time horizon. So it is called "integral" quadratic constraints. For discrete-time operators, we just sum things up. We require (7.5) to hold for any N. In controls literature, this type of constraints are "hard" IQCs. We will briefly talk about "soft" IQCs in some future lecture when we discuss the KYP lemma. For now, we focus on hard IQCs that are in the form of (7.5). Hard IQCs can be directly incorporated into the dissipation inequality framework. Typically hard IQCs lead to a supply rate condition $\sum_{k=0}^{N} S(\xi_k, w_k) \leq 0$. Suppose we have constructed a dissipation inequality $V(\xi_{k+1}) - V(\xi_k) \leq S(\xi_k, w_k)$. Now we do not have $S \leq 0$ for all k. However, we can first sum up the dissipation inequality from $k = 0$ to N to get $V(\xi_{N+1}) \leq V(\xi_0) + \sum_{k=0}^{N} S(\xi_k, w_k)$. Now using the new supply rate condition $\sum_{k=0}^{N} S(\xi_k, w_k) \leq 0$, we obtain $V(\xi_{N+1}) \leq V(\xi_0)$. Hence the internal energy is bounded. The physical interpretation is that as long as the total energy supplied to the system (which is equal to $\sum_{k=0}^{N} S(\xi_k, w_k)$) is non-positive, the internal energy is not going to be larger than the initial energy. We will talk about how to use IQCs for convergence rate analysis later. There is a routine for that.

IQCs are more general than pointwise quadratic constraints. Whenever we have the pointwise quadratic constraint (7.1), we immediately have an IQC in the form of (7.5) by summing the constraints from $k = 0$ to N. The reverse direction is not always true. When Δ has dynamics and memory, it is very common that we will only be able to construct IQCs.

Example: A general version of small gain bound. Consider a general causal operator Δ. A general version of the small gain bound enforces the following inequality for the input/output pair of Δ:

$$\sum_{k=0}^{N} \|w_k - w^*\|^2 \leq L^2 \sum_{k=0}^{N} \|v_k - v^*\|^2.$$

This bound is equivalent to the following IQC:

$$\sum_{k=0}^{N} \begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix}^\mathsf{T} \begin{bmatrix} -L^2 I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} v_k - v^* \\ w_k - w^* \end{bmatrix} \leq 0. \tag{7.6}$$

In general, when Δ is the so-called "bounded" operator, we will always have the above small gain IQC. For example, if Δ is an unknown stable LTI system whose $\mathcal{H}_\infty$ norm is L, then we will not have a pointwise small gain bound but (7.6) still holds with $v^* = w^* = 0$. You can verify a similar fact when Δ is a time-varying delay. In many situations, even for static Δ, we can construct useful IQCs to complement the use of pointwise constraints. We will see more examples in future lectures.