Massachusetts Institute of Technology

Department of Electrical Engineering and Computer Science

6.245: MULTIVARIABLE CONTROL SYSTEMS

by A. Megretski

Hankel Optimal Model Order Reduction[1]

This lecture covers both the theory and an algorithmic side of Hankel optimal model order reduction.

9.1 Problem Formulation and Main Results

This section formulates the main theoretical statements of Hankel optimal model reduction, including the famous Adamyan-Arov-Krein (AAK) theorem.

9.1.1 The Optimization Setup

Let $G = G(s)$ be a matrix-valued function bounded on the $j\omega$-axis. The task of Hankel optimal model reduction of G calls for finding a stable LTI system $\hat{G}$ of order less than a given positive integer m, such that the Hankel norm $\|\Delta\|_H$ of the difference $\Delta = G - \hat{G}$ is minimal. Remember that Hankel norm of an LTI system with transfer matrix $\Delta = \Delta(s)$, input w, and output v, is defined as the L2 gain of the associated Hankel operator $\mathcal{H}_\Delta$, i.e. as the maximum of the "future output energy integral"

$$\left(\int_0^\infty |v(t)|^2 dt\right)^{1/2}$$

subject to the constraints

$$w(t) = 0 \text{ for } t \geq 0, \quad \int_{-\infty}^0 |w(t)|^2 dt \leq 1.$$

---

[1] Version of March 29, 2004

While the standard numerical algorithms of Hankel optimal model reduction will be defined here for the case when G is a stable causal finite order LTI system defined by a controllable and observable state space model

$$\dot{x} = Ax + Bw, \;\; y = Cx + Dw, \tag{9.1}$$

some useful insight can be obtained by studying general statements when $G = G(s)$ is not necessarily a rational transfer matrix..

9.1.2 The i-th largest singular value of a Hankel operator

Let $G = G(s)$ be a matrix-valued function bounded on the $j\omega$-axis.

For an integer $k > 0$ let us say that $\sigma = \sigma_k$ is the k-th singular number of $\mathcal{H}_G$ if σ_k is the minimum of the L2 gain of $\Delta - \mathcal{H}_G$ over the set of all linear transformations Δ of rank less than k (this definition works for arbitrary linear transformations, not only for Hankel operators).

When G is defined by a minimal state space model (9.1) with m states and a Hurwitz matrix A, $\sigma_i = 0$ for $i > m$, and the first m largest singular numbers s_i are square roots of the corresponding eigenvalues of $W_o W_c$, where W_o, W_c are the controllability and observability Grammians of (9.1). For some non-rational transfer matrices, an analytical or numerical calculation of σ_i may be possible. For example, the i-th largest singular number of $\mathcal{H}_G$, where $G(s) = \exp(-s)$, equals 1 for all positive i.

9.1.3 The AAK Theorem

The famous Adamyan-Arov-Krein theorem provides both a theoretical insight and (taking a constructive proof into account) an explicit algorithm for finding Hankel optimal reduced models.

Theorem 9.1 *Let $G = G(s)$ be a matrix-valued function bounded on the $j\omega$-axis. Let $\sigma_1 \geq \sigma_2 \geq \dots \sigma_m \geq 0$ be the m largest singular values of $\mathcal{H}_G$. Then σ_m is the minimum of $\|G - \hat{G}\|$ over the set of all stable systems $\hat{G}$ of order less than m.*

In other words, approximating Hankel operators by general linear transformations of rank less than m cannot be done better (in terms of the minimal L2 gain of the error) than approximating it by Hankel operators of rank less than m.

The proof of the theorem, to be given in the next section for the case of a rational transfer matrix $G = G(s)$, is constructive, and provides a simple algorithm for calculating the Hankel optimal reduced model.

9.1.4 H-Infinity quality Hankel optimal reduced models

It is well established by numerical experiments that Hankel optimal reduced models usually offer very high H-Infinity quality of model reduction. A somewhat conservative description of this effect is given by the following extension of the AAK theorem.

Theorem 9.2 *Let $G = G(s)$ be a stable rational function. Let $\sigma_1 \geq \sigma_2 \geq \dots \sigma_m > 0$ be the m largest singular values of $\mathcal{H}_G$. Let $\sigma_m > \tilde{\sigma}_{m+1} > \tilde{\sigma}_{m+2} > \cdots > \tilde{\sigma}_r = 0$ be the ordered sequence of the remaining singular values of $\mathcal{H}_G$, each value taken once, without repetition. Then there exists a Hankel optimal reduced model $\hat{G}$ of order less than m such that*

$$\|G - \hat{G}\|_\infty \leq \sigma_m + \sum_{k=m+1}^{r} \tilde{\sigma}_k.$$

Just as in the case of the basic AAK theorem, the proof of Theorem 9.2 is constructive, and hence provides an explicit construction of the reduced model with the proven properties. In practice, the actual H-Infinity norm of model reduction error is much smaller.

It is important to remember that the Hankel optimal reduced model is never unique (at least, the "D" terms do not have any effect on the Hankel norm, and hence can be modified arbitrarily). The proven H-Infinity model reduction error bound is guaranteed only for a specially selected Hankel optimal reduced model.

9.2 Proof of the AAK theorem

The fact that σ_m is a *lower bound* for the Hankel norm model reduction error follows from the definition of σ_m and from the fact that rank of the Hankel operator associated with a stable system equals its order. This section contains a rather explicit construction of a reduced model of order less than m for which the Hankel norm of the approximation error equals σ_m. To avoid inessential mathematical complications, we consider only the case when G is defined by a minimal (controllable and observable) finite dimensional state space model

$$\dot{x} = Ax + Bw, \;\; y = Cx + Dw \tag{9.2}$$

with a Hurwitz matrix A.

The starting point of the proof is the fact that, according to the explicit formulae for singular value decomposition of Hankel operators based of the controllability and observability Grammians W_c, W_o of system (9.2), the matrix $W_o - \sigma_m^2 W_c^{-1}$ has not more than $m-1$ positive eigenvalues. The proofs is similar to the one given in the derivation of H-Infinity suboptimal control, and also uses the generalized Parrot's theorem in combination with a modified version of the KYP lemma.

9.2.1 Upper bounds for Hankel norms

The following simple observation is frequently helpful when workin with Hankel operators.

Theorem 9.3 *If rational transfer matrices G_+, G_- of same dimensions are such that all poles of G_+ have negative real part, and all poles of G_- have positive real part, then*

$$\|G_+\|_H \leq \|G_+ + G_-\|_\infty.$$

Since the Hankel operator associates with a transfer matrix $G = G_+ + G_-$ is defined by the stable part G_+ of G, and the Hankel norm never exceeds L-Infinity norm, the statement is true. In fact, Theorem 9.3 is the "simple" side of the so-called Nehari theorem, which claims that $\|G_+\|_H$ is the *minimum* of $\|G_+ + G_-\|_\infty$ over the set of all anti-stable G_-. The non-trivial part of the Nehari theorem will be a by-product of the proof of the AAK theorem.

Example 9.1 Adding $G_- = -1/2$ to $G_+ = 1/(s+1)$ yields $G = (1-s)/(2+2s)$ (Infinity-norm $\|G\|_\infty = 0.5$ equals the Hankel norm of $1/(s+1)$).

9.2.2 KYP lemma for L-Infinity norm approximation

A simple but important observation given by the KYP lemma is that a "certificate" of an L-Infinity bound $\|H\|_\infty \le \gamma$ for a given transfer matrix $H(s) = d + c(sI - a)^{-1}b$ is delivered by a symmetric matrix $p = p'$ such that

$$\gamma^2|w|^2 - |c\bar{x} + dw|^2 - 2\bar{x}'p(a\bar{x} + bw) \ge 0 \quad \forall\ \bar{x}, w. \tag{9.3}$$

Note that a does not have to be a Hurwitz matrix here.

When system (9.2) is approximated by system G_r (not necessarily stable) with state space model

$$\dot{x}_r = A_r x_r + B_r w, \quad y_r = C_r x_r + D_r w, \tag{9.4}$$

the "approximation error dynamics" system with input w and output $\delta = y - y_r$ has state space model

$$\dot{\bar{x}} = a\bar{x} + bw, \quad \delta = c\bar{x} + dw,$$

where

$$\bar{x} = \begin{bmatrix} x \\ x_r \end{bmatrix}, \; a\bar{x} + bw = \begin{bmatrix} Ax + Bw \\ A_r x_r + B_r w \end{bmatrix}, \; c\bar{x} + dw = Cx + Dw - C_r x_r - D_r w. \tag{9.5}$$

Let $\Delta = \Delta(s)$ denote the transfer matrix from w to δ. According to the KYP lemma, the inequality $\|\Delta\|_\infty \le \gamma$ can be established by finding a symmetric matrix

$$p = p' = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \tag{9.6}$$

such that (9.3) holds.

In view of Theorem 9.3, for Hankel model order reduction it is important to keep track of the order of the stable part of G_r. Tis is made possible by the following observation.

Theorem 9.4 *If (9.3) holds for a, b, c, d, p defined by (9.5),(9.6), and p_{22} has less than m positive eigenvalues then the order of the stable part of G_r is less than m.*

Proof Let $V = V_+ \dot{+} V_-$ be the direct sum decomposition of the state space $\{x_r\}$ into the *stable observable* subspace V_+ of A_r with respect to C_r, and the compementary A_r-invariant subspace V_-. In a system of coordinates associated with this decomposition, matrices C_r, A_r, p_{22} have the block form

$$C_r = \begin{bmatrix} c_+ & c_- \end{bmatrix}, \quad A_r = \begin{bmatrix} a_+ & 0 \\ 0 & a_- \end{bmatrix}, \quad p_{22} = \begin{bmatrix} p_{++} & p_{+-} \\ p_{-+} & p_{--} \end{bmatrix},$$

where the pair (c_+, a_+) is observable. Substituting $w = 0$ into (9.3) yields

$$pa + a'p \leq -c'c,$$

which, in particular, implies

$$p_{++}a_+ + a'_+p_{++} \leq -c'_+c_+.$$

Hence $p_{++} > 0$ and the number of positive eigenvalues of p_{22} is at least as large as the dimension of a_+. Finally, note that the dimension of a_+ is not smaller than the order of the stable part of G_r. ■

9.2.3 Hankel optimal reduced models via Parrot's Theorem

The observations made in the previous two subsections suggest the following approach to constucting a reduced model $\hat{G}$ of system G given by (9.2): simply find a matrix $p = p'$ in (9.6) such that p_{22} has less than $n + m$ positive eigenvalues, and (9.3) holds for a, b, c, d defined by (9.5) with some A_r, B_r, C_r, D_r. Then $\hat{G}$, defined as the stable part of $D_r + C_r(sI - A_r)^{-1}B_r$, will satisfy $\|G - \hat{G}\|_H \leq \gamma$.

Note that, once p is fixed, the existence of A_r, B_r, C_r, D_r satisfying the requirements can be checked via the generalized Parrot's theorem (same as used in the derivation of H-Infinity suboptimal controllers).

Theorem 9.5 *Let* $\sigma : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^k \to \mathbf{R}$ *be a quadratic form which is concave with respect to its second argument, i.e.*

$$\sigma(0, g, 0) \leq 0 \quad \forall\, g \in \mathbf{R}^m. \tag{9.7}$$

Then an m*-by-*k *matrix* L *such that*

$$\sigma(f, Lh, h) \geq 0 \quad \forall\, f \in \mathbf{R}^n, h \in \mathbf{R}^k \tag{9.8}$$

exists if and only if the following two conditions are satisfied:

(a) for every $h \in \mathbf{R}^k$ *there exists* $g \in \mathbf{R}^m$ *such that*

$$\inf_{f \in \mathbf{R}^n} \sigma(f, g, h) > -\infty; \tag{9.9}$$

(b) the inequality

$$\sup_{g \in \mathbf{R}^m} \sigma(f, g, h) \geq 0 \tag{9.10}$$

holds for all $f \in \mathbf{R}^n$, $h \in \mathbf{R}^k$.

In application to Hankel optimal model reduction, set

$$f = x,\ g = \begin{bmatrix} \theta \\ y_r \end{bmatrix},\ h = \begin{bmatrix} x_r \\ w \end{bmatrix},\ L = \begin{bmatrix} A_r & B_r \\ C_r & D_r \end{bmatrix},$$

$$\sigma(f, g, h) = -2 \begin{bmatrix} x \\ x_r \end{bmatrix}' p \begin{bmatrix} Ax + Bw \\ \theta \end{bmatrix} + \gamma^2 |w|^2 - |Cx + Dw - y_r|^2. \tag{9.11}$$

Note that σ is concave with respect to g.

It turns out that one convenient selection for p is

$$p = \begin{bmatrix} W_o & W_o - \gamma^2 W_c^{-1} \\ W_o - \gamma^2 W_c^{-1} & W_o - \gamma^2 W_c^{-1} \end{bmatrix}. \tag{9.12}$$

While formally there is no need to explain this choice (one just has to verify that conditions (a),(b) of Theorem 9.5 are satisfied), there is a clear line of reasoning here. To come to this particular choice of p, note first that (9.9) implies $\sigma(f, 0, 0) \geq 0$, which means

$$p_{11} A + A' p_{11} \leq -C'C$$

for σ defined by (9.11). Hence

$$p_{11} \geq W_o.$$

Similarly, under the simplifyin assumption that p is invertible, (9.10) means

$$A q_{11} + q_{11} A' \leq -\gamma^{-2} B B',$$

where q_{11} is the upper left corner of p^{-1}. Hence

$$q_{11} \geq \gamma^{-2} W_c.$$

Since

$$q_{11}^{-1} = p_{11} - p_{12} p_{22}^{-1} p_{21},$$

we have

$$p_{12} p_{22}^{-1} p_{21} = p_{11} - \gamma^2 q_{11}^{-1}.$$

Since our desire is to minimize the number of positive eigenvalues of p_{22}, it is natural to use the minimal possible values of p_{11} and q_{11}, which suggests using

$$p_{11} = W_o, \quad q_{11} = \gamma^{-2} W_c,\ p_{12} = p_{21} = p_{22} = W_o - \gamma^2 W_c^{-1}.$$

With the proposed selection of p, the quadratic form σ can be represented as

$$\sigma(f,g,h) = -2x'[W_oBw + A'(W_o - \gamma^2 W_c^{-1})x_r + (W_o - \gamma^2 W_c^{-1})\theta - C'y_r] + \tilde{\sigma}(g,h).$$

Hence condition (a) can be satisfied if the equation

$$W_oBw + A'(W_o - \gamma^2 W_c^{-1})x_r + (W_o - \gamma^2 W_c^{-1})\theta - C'y_r = 0$$

has a solution (θ, y_r) for every pair (w, x_r). In order to prove this, it is sufficient to show that every vector ψ such that

$$\psi' C' = 0, \quad \psi'(W_o - \gamma^2 W_c^{-1}) = 0 \tag{9.13}$$

also satisfies

$$\psi' W_o B = 0, \quad \psi' A'(W_o - \gamma^2 W_c^{-1}) = 0. \tag{9.14}$$

Note that by the definition,

$$(W_o - \gamma^2 W_c^{-1})A + A'(W_o - \gamma^2 W_c^{-1}) + C'C = \gamma^2 W_c^{-1} BB' W_c^{-1}.$$

Multiplying this by ψ' on the left and ψ on the right and using (9.13) yields $\psi' B' W_c^{-1} = 0$ and hence (9.14) follows.

Similarly, the quadratic form σ can be represented as

$$\sigma(f,g,h) = (x+x_r)'(W_o-\gamma^2 W_c^{-1})(Ax+Bw-\theta)+2\gamma^2 x' W_c^{-1}(Ax+Bw)+\gamma^2|w|^2-|Cx-y_r|^2,$$

which is unbounded with respect to θ if $(x + x_r)'(W_o - \gamma^2 W_c^{-1}) \neq 0$ and is made non-negative by selecting $y_r = -Cx_r$ otherwise.