

Engr210a Lecture 11: Hankel Operators and Model Reduction

- Stability of balanced truncation
- Inner functions.
- Error-bounds for balanced truncation.
- Examples.
- Singular perturbation.
- Optimal Hankel-norm approximation.
- Optimal induced-norm model reduction.

Balanced truncation

Given G of order n , we wish to find a reduced-order model of order $r < n$. Suppose $D = 0$, and A, B, C is a balanced realization for G . Partition matrices A, B, C as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \quad C = [C_1 \ C_2]$$

where $A_{11} \in \mathbb{R}^{r \times r}$. The reduced order model will be

$$\hat{G}_r = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & 0 \end{array} \right]$$

This reduced-order model is called a *balanced truncation* of G .

Notes

- Assume $\sigma_r > \sigma_{r+1}$. That is, these singular values must not be equal.
- We will show that G_r is stable and balanced, and derive an upper bound on the modeling error

$$\|G - G_r\|$$

- The method of *truncation* is an example of a *Galerkin projection* of the differential equations onto a particular basis; the basis we are using is that spanned by the r most controllable and observable states.

Stability of Balanced Truncation

Suppose $(A, B, C, 0)$ is a minimal balanced realization for the stable system G with state-dimension n . Suppose $(A_{11}, B_1, C_1, 0)$ is the balanced truncation of G with state-dimension $r < n$, and assume $\sigma_r > \sigma_{r+1}$. Then

- (i) A_{11} is Hurwitz
- (ii) $(A_{11}, B_1, C_1, 0)$ is balanced.

Proof: The Lyapunov equations for controllability and observability are

$$\begin{bmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} + \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} C_1^* \\ C_2^* \end{bmatrix} \begin{bmatrix} C_1 & C_2 \end{bmatrix} = 0$$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} + \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{bmatrix} B_1^* & B_2^* \end{bmatrix} = 0$$

The (1,1) blocks of this matrix equation are

$$A_{11}^* \Sigma_1 + \Sigma_1 A_{11} + C_1^* C_1 = 0$$

$$A_{11} \Sigma_1 + \Sigma_1 A_{11}^* + B_1 B_1^* = 0$$

Hence if A_{11} is Hurwitz, then $(A_{11}, B_1, C_1, 0)$ is balanced, with Hankel singular values $\sigma_1, \dots, \sigma_r$. This proves part (ii) if part (i) holds.

Proof of Stability of Balanced Truncation, continued

- We wish to prove that A_{11} is Hurwitz.
- Suppose $A_{11}v = \lambda v$. Then

$$\begin{aligned} v^*(A_{11}^*\Sigma_1 + \Sigma_1 A_{11} + C_1^*C_1)v &= 0 \\ \implies 2\operatorname{Re}(\lambda) &= -\frac{v^*C_1^*C_1v}{v^*\Sigma_1v} \leq 0 \end{aligned}$$

- So all we need to show is that A_{11} cannot have any imaginary eigenvalues.
- We will prove this by contradiction. Suppose A_{11} has an imaginary eigenvalue, and let $V = [v_1 \dots v_p] \in \mathbb{R}^{n \times p}$ satisfy

$$\operatorname{image}(V) = \ker(j\omega I - A_{11}) \implies A_{11}V = j\omega V$$

Then the above argument shows $v_i^*C_1^*C_1v_i = 0$ for $i = 1, \dots, p$, which implies $C_1V = 0$. Then

$$\begin{aligned} (A_{11}^*\Sigma_1 + \Sigma_1 A_{11} + C_1^*C_1)V &= 0 \\ \implies A_{11}^*\Sigma_1V &= -j\omega\Sigma_1V \end{aligned}$$

Proof of Stability of Balanced Truncation, continued

- So far, we have $A_{11}V = j\omega V$, and $A_{11}^*\Sigma_1 V = -j\omega\Sigma_1 V$.

- We know

$$V^*\Sigma_1(A_{11}\Sigma_1 + \Sigma_1 A_{11}^* + B_1 B_1^*)\Sigma_1 V = 0$$

$$\implies j\omega V^*\Sigma_1^3 V - j\omega V^*\Sigma_1^3 V + V^*\Sigma_1 B_1 B_1^*\Sigma_1 V = 0$$

$$\implies B_1^*\Sigma_1 V = 0$$

- Hence

$$(A_{11}\Sigma_1 + \Sigma_1 A_{11}^* + B_1 B_1^*)\Sigma_1 V = 0$$

$$\implies (A_{11} - j\omega I)\Sigma_1^2 V = 0$$

$$\implies \text{image}(\Sigma_1^2 V) \subseteq \ker(j\omega I - A_{11}) = \text{image}(V)$$

- Hence V is an *invariant subspace* of Σ_1^2 . Hence there exists $q \in \text{image}(V)$ such that

$$\Sigma_1^2 q = \sigma_i^2 q$$

for some i with $1 \leq i \leq r$.

- Since $q \in \text{image}(V)$, we have $A_{11}q = j\omega q$. We will show that

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} q \\ 0 \end{bmatrix} = j\omega \begin{bmatrix} q \\ 0 \end{bmatrix}$$

Proof of Stability of Balanced Truncation, continued

- The Lyapunov equations for controllability and observability give

$$\left(\begin{bmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} + \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} C_1^* \\ C_2^* \end{bmatrix} [C_1 \ C_2] \right) \begin{bmatrix} q \\ 0 \end{bmatrix} = 0$$

$$\left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} + \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} [B_1^* \ B_2^*] \right) \begin{bmatrix} \Sigma_1 q \\ 0 \end{bmatrix} = 0$$

- $q \in \text{image}(V)$ implies $C_1 q = 0$ and $B_1^* \Sigma_1 q = 0$. The above equations then give

$$A_{12}^* \Sigma_1 q + \Sigma_2 A_{21} q = 0$$

$$A_{21} \sigma_i^2 q + \Sigma_2 A_{12}^* \Sigma_1 q = 0$$

which implies $\Sigma_2^2 A_{21} q = \sigma_i^2 A_{21} q$. But we know $1 \leq i \leq r$ and Σ_1 and Σ_2 have no common eigenvalues, so $A_{21} q = 0$.

- Hence

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} q \\ 0 \end{bmatrix} = j\omega \begin{bmatrix} q \\ 0 \end{bmatrix}$$

which means that A has an imaginary eigenvalue, which is a contradiction.

- Note that if Σ_1 and Σ_2 have common eigenvalues ($\sigma_r = \sigma_{r+1}$), then A_{11} can be unstable.

Inner functions

A transfer function $\hat{U} \in H_\infty$ which satisfies

$$(\hat{U}(j\omega))^* \hat{U}(j\omega) = I \quad \text{for all } \omega \in \mathbb{R}$$

is called an *inner function*.

Notes

- If $\hat{G}(j\omega) = (\hat{U}(j\omega))^*$ then $M_{\hat{G}} = M_{\hat{U}}^*$.
- $M_{\hat{U}}$ is an isometry, since $\langle M_{\hat{U}}x, M_{\hat{U}}x \rangle = \langle x, M_{\hat{U}}^* M_{\hat{U}}x \rangle = \langle x, x \rangle$
- $\bar{\sigma}(\hat{U}(j\omega)) = \bar{\sigma}\left((\hat{U}(j\omega))^* \hat{U}(j\omega)\right)^{\frac{1}{2}} = 1$ for all $\omega \in \mathbb{R}$

The transfer function has unit gain at every frequency. For this reason, inner functions are also called *all-pass* functions.

Para-hermitian conjugate

Suppose \hat{U} has realization (A, B, C, D) . Then

$$\hat{U}^\sim = \left[\begin{array}{c|c} -A^* & -C^* \\ \hline B^* & D^* \end{array} \right]$$

is called the *para-hermitian conjugate* of \hat{U} .

Notes

- $\hat{U}^\sim(j\omega) = (\hat{U}(j\omega))^*$ for all $\omega \in \mathbb{R}$.
- If $\hat{U} \in H_\infty$, then \hat{U}^\sim is analytic on the closed left-half plane. The matrix $-A^*$ is unstable if A is stable.

State-space test for inner functions

Suppose $\hat{G} \in H_\infty$ has a realization (A, B, C, D) where A is Hurwitz. Let Y_o be the observability gramian for (A, C) . Then

$$C^*D + Y_oB = 0 \quad \implies \quad (\hat{U}(s))^* \hat{U}(s) = D^*D \quad \text{for all } s \in \mathbb{C}$$

Proof

A realization for $\hat{U} \sim \hat{U}$ is

$$\hat{U} \sim \hat{U} = \left[\begin{array}{cc|c} -A^* & -C^*C & -C^*D \\ 0 & A & B \\ \hline B^* & D^*C & D^*D \end{array} \right]$$

Changing state-space coordinates under transformation $T = \begin{bmatrix} I & -Y_o \\ 0 & I \end{bmatrix}$ gives

$$\hat{U} \sim \hat{U} = \left[\begin{array}{cc|c} -A^* & -(A^*Y_o + Y_oA + C^*C) & -(C^*D + Y_oB) \\ 0 & A & B \\ \hline B^* & D^*C + B^*Y_o & D^*D \end{array} \right] = \left[\begin{array}{cc|c} -A^* & 0 & 0 \\ 0 & A & B \\ \hline B^* & 0 & D^*D \end{array} \right]$$

All states in this realization are either uncontrollable or unobservable.

Error-bounds for balanced truncation

Assume

- $(A, B, C, 0)$ is a balanced realization for G , with order n .
This implies A is Hurwitz, and $(A, B, C, 0)$ is minimal.
- G_r is the balanced truncation of G , with realization $(A_{11}, B_1, C_1, 0)$ and order r .
- The Hankel singular values of G satisfy $\sigma_i = \sigma_{r+1}$ for $i = r+1, \dots, n$ and $\sigma_r > \sigma_{r+1}$.

That is,

$$Y_o = X_c = \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \sigma_{r+1} I \end{bmatrix} \quad \text{where} \quad \Sigma_1 = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \dots & \\ & & & \sigma_r \end{bmatrix}$$

Theorem

The induced-norm error between G and G_r satisfies

$$\|G - G_r\| \leq 2\sigma_{r+1}$$

Proof

- Let $F = G - G_r$ be the error system. F has realization

$$\hat{F} = \left[\begin{array}{ccc|c} A_{11} & 0 & 0 & B_1 \\ 0 & A_{11} & A_{12} & B_1 \\ 0 & A_{21} & A_{22} & B_2 \\ \hline -C_1 & C_1 & C_2 & 0 \end{array} \right]$$

By the previous result, A_{11} is Hurwitz, so F is stable also.

- The idea of the proof is to add inputs and outputs to the system to create a new transfer function

$$\hat{E}(s) = \begin{bmatrix} \hat{F}(s) & \hat{E}_{12}(s) \\ \hat{E}_{21}(s) & \hat{E}_{22}(s) \end{bmatrix}$$

which is inner. This is called an *all-pass dilation* of F .

- Clearly $\|F\| \leq \|E\|$.

Proof, continued

- Using state coordinate transformation

$$T = \begin{bmatrix} I & I & 0 \\ I & -I & 0 \\ 0 & 0 & I \end{bmatrix} \quad \text{results in} \quad \hat{F} = \left[\begin{array}{ccc|c} A_{11} & 0 & A_{12}/2 & B_1 \\ 0 & A_{11} & -A_{12}/2 & 0 \\ A_{21} & -A_{21} & A_{22} & B_2 \\ \hline 0 & -2C_1 & C_2 & 0 \end{array} \right]$$

- The all-pass dilation is

$$\hat{E}(s) = \left[\begin{array}{ccc|cc} A_{11} & 0 & A_{12}/2 & B_1 & 0 \\ 0 & A_{11} & -A_{12}/2 & 0 & \sigma_{r+1}\Sigma_1^{-1}C_1^* \\ A_{21} & -A_{21} & A_{22} & B_2 & -C_2^* \\ \hline 0 & -2C_1 & C_2 & 0 & 2\sigma_{r+1}I \\ -2\sigma_{r+1}B_1^*\Sigma_1^{-1} & 0 & -B_2^* & 2\sigma_{r+1}I & 0 \end{array} \right]$$

which has observability gramian $\bar{Y}_o = \begin{bmatrix} 4\sigma_{r+1}^2\Sigma_1^{-1} & 0 & 0 \\ 0 & 4\Sigma_1 & 0 \\ 0 & 0 & 2\sigma_{r+1}I \end{bmatrix}$

- One can verify this is inner, with $(\hat{E}(j\omega))^* \hat{E}(j\omega) = 4\sigma_{r+1}^2$.
- Hence $\|G - G_r\| = \|F\| \leq \|E\| = 2\sigma_{r+1}$.

General error-bounds for balanced truncation

Assume

- $(A, B, C, 0)$ is a balanced realization for G , with order n .
This implies A is Hurwitz, and $(A, B, C, 0)$ is minimal.
- G_r is the balanced truncation of G , with realization $(A_{11}, B_1, C_1, 0)$ and order r .
- The Hankel singular values of G satisfy $\sigma_r > \sigma_{r+1}$.

Let $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$ where

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \dots & \\ & & & \sigma_r \end{bmatrix} \quad \Sigma_2 = \begin{bmatrix} \sigma_1^t I & & & \\ & \sigma_2^t I & & \\ & & \dots & \\ & & & \sigma_k^t I \end{bmatrix}$$

and the σ^t are distinct.

Theorem

The induced-norm error between G and G_r satisfies

$$\|G - G_r\| \leq 2(\sigma_1^t + \dots + \sigma_k^t)$$

Theorem

The induced-norm error between G and G_r satisfies

$$\|G - G_r\| \leq 2(\sigma_1^t + \cdots + \sigma_k^t)$$

Proof

- Truncate the states corresponding to $\sigma_k^t, \sigma_{k-1}^t, \dots, \sigma_1^t$.
- Let $G^{(k)} = G$.
- Let $G^{(i-1)}$ be the balanced truncation of $G^{(i)}$, removing states corresponding to σ_i^t . Then $G^{(0)} = G_r$.
- Note that $G^{(i)}$ is also a balanced truncation of G .
- Applying the previous result at each stage gives

$$\begin{aligned} \|G - G_r\| &= \|(G^{(k)} - G^{(k-1)}) + (G^{(k-1)} - G^{(k-2)}) + \cdots + (G^{(1)} - G^{(0)})\| \\ &\leq \|G^{(k)} - G^{(k-1)}\| + \|G^{(k-1)} - G^{(k-2)}\| + \cdots + \|G^{(1)} - G^{(0)}\| \\ &\leq 2(\sigma_1^t + \cdots + \sigma_k^t) \end{aligned}$$

Balanced truncation

The induced-norm error between G and G_r satisfies

$$\|G - G_r\| \leq 2(\sigma_1^t + \cdots + \sigma_k^t)$$

Notes

- This formula is known as the *twice-the-sum-of-the-tail* formula.
- Applying it to the zero order truncation gives

$$\|G\| \leq 2(\sigma_1 + \cdots + \sigma_n) \quad \text{excluding multiplicities}$$

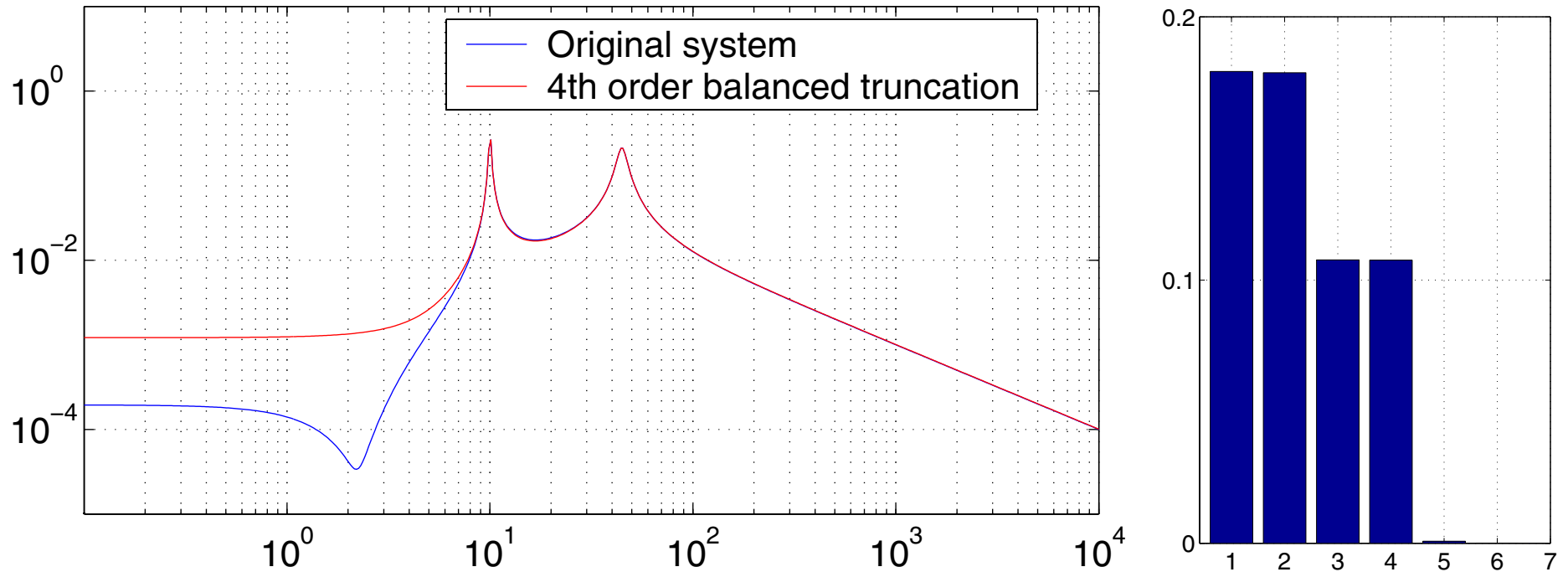
- This is an *upper bound* on the error. In general, the balanced-truncation error may be much less than this, but cannot be less than σ_{r+1} .
- This result was first proved by Dale Enns, in his 1984 Ph.D. thesis *Model reduction for control system design* at Stanford. It was also independently proved by Keith Glover in Cambridge in 1984.
- Since then, there have been many developments, including extensions for nonlinear and uncertain systems, PDEs, time-varying systems, etc.

Example

$$\hat{G}(s) = \frac{(s + 10)(s - 5)(s^2 + 2s + 5)(s^2 - 0.5s + 5)}{(s + 4)(s^2 + 4s + 8)(s^2 + 0.2s + 100)(s^2 + 5s + 2000)},$$

$$\Sigma = \text{diag}(0.1793, 0.1789, 0.1077, 0.1076, 0.00076, 0.00008, 0.00003)$$

transfer function magnitude versus frequency



$$\|G - G_4\| = 0.0014$$

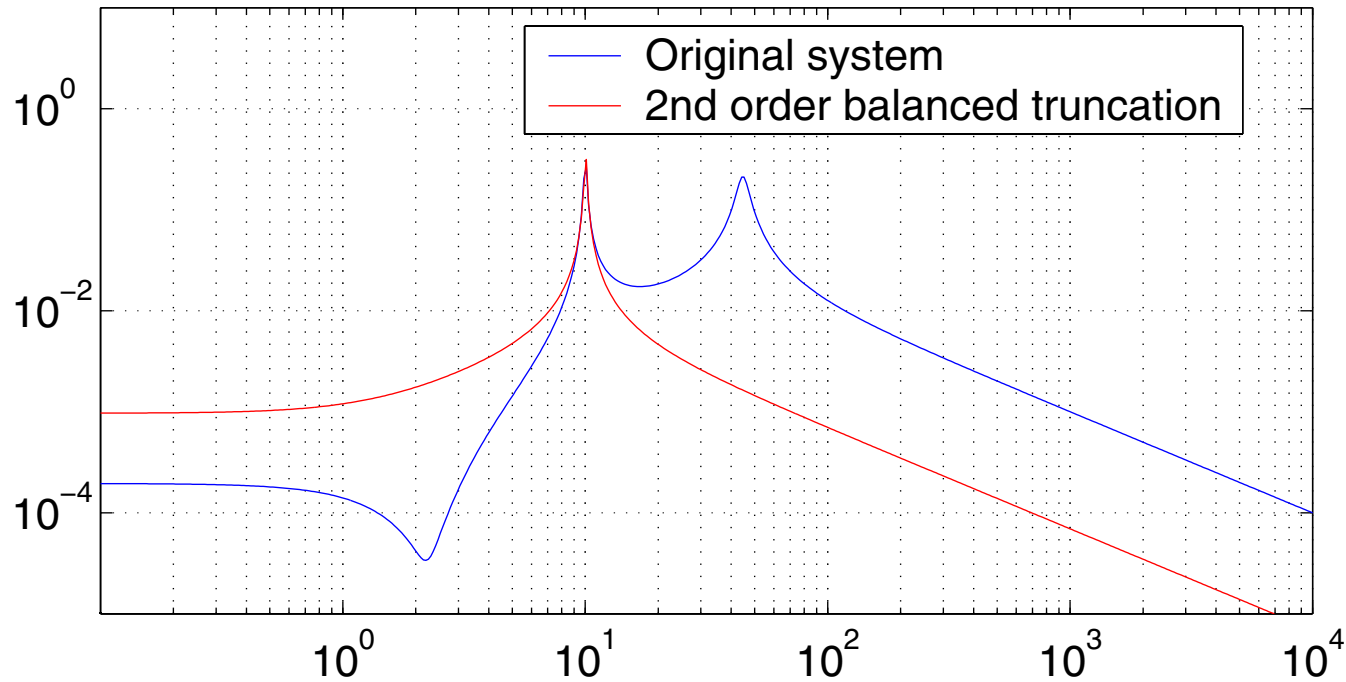
$$0.00076 = \sigma_5 \leq \|G - G_4\| \leq 2(\sigma_5 + \sigma_6 + \sigma_7) = 0.0017$$

Example

$$\hat{G}(s) = \frac{(s + 10)(s - 5)(s^2 + 2s + 5)(s^2 - 0.5s + 5)}{(s + 4)(s^2 + 4s + 8)(s^2 + 0.2s + 100)(s^2 + 5s + 2000)},$$

$$\Sigma = \text{diag}(0.1793, 0.1789, 0.1077, 0.1076, 0.00076, 0.00008, 0.00003)$$

transfer function magnitude versus frequency



$$\|G - G_2\| = 0.2153$$

$$0.1077 = \sigma_3 \leq \|G - G_2\| \leq 2(\sigma_3 + \sigma_4\sigma_5 + \sigma_6 + \sigma_7) = 0.4322$$

Balanced singular perturbation

Given state-space system

$$\dot{x}_1(t) = A_{11}x_1(t) + A_{12}x_2(t) + B_1u(t)$$

$$\dot{x}_2(t) = A_{21}x_1(t) + A_{22}x_2(t) + B_2u(t)$$

$$y(t) = C_1x_1(t) + C_2x_2(t) + Du(t)$$

The truncation method gives zero error as frequency tends to ∞ . The error at zero-frequency associated with state-space truncation is

$$\hat{G}(0) - \hat{G}_r(0) = CA^{-1}B - C_1A_{11}^{-1}B_1$$

An alternative is *singular perturbation*. Construct $\hat{H}(s) = \hat{G}(s^{-1})$, and apply truncation to construct \hat{H}_r . Then let $\hat{G}_r(s) = \hat{H}_r(s^{-1})$. This gives zero error at zero frequency. State-space formula are

$$\hat{G}_r = \left[\begin{array}{c|c} A_{11} - A_{12}A_{22}^{-1}A_{21} & B_1 - A_{12}A_{22}^{-1}B_2 \\ \hline C_1 - C_2A_{22}^{-1}A_{21} & D - C_2A_{22}^{-1}B_2 \end{array} \right]$$

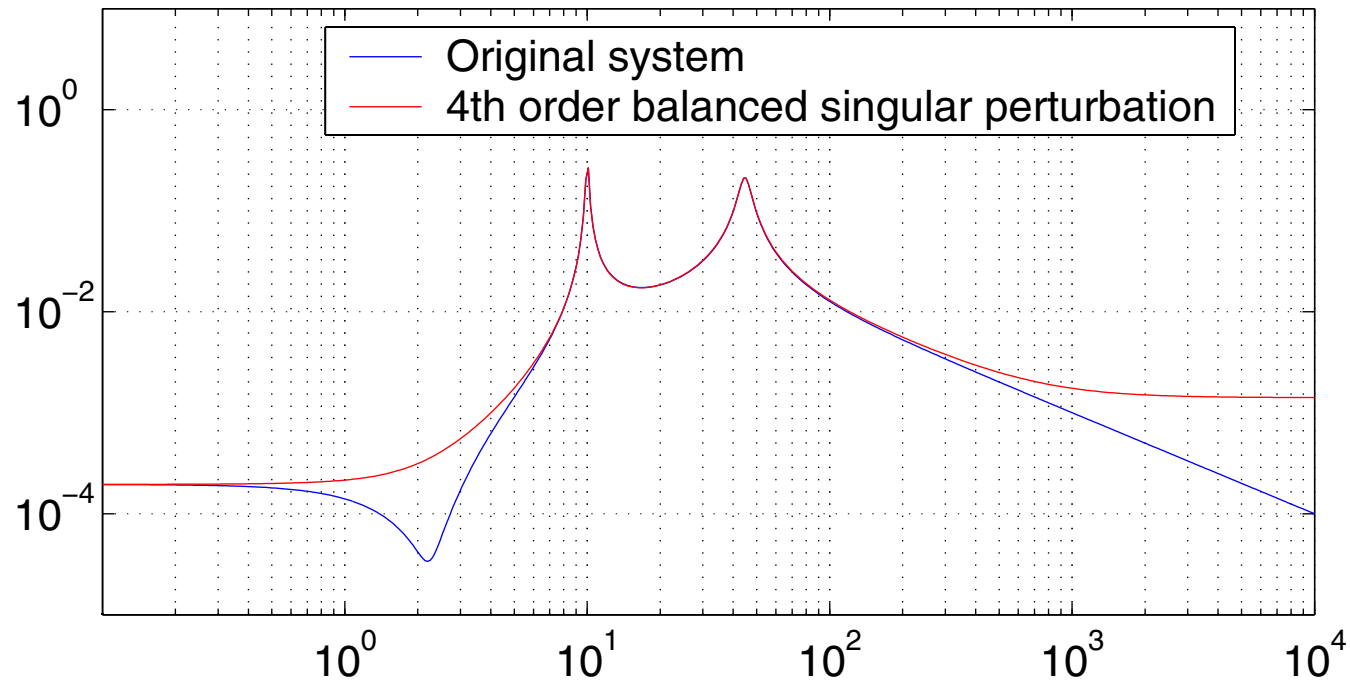
When combined with balancing, this is called *balanced singular perturbation*. The transformation maps the left-half-plane to itself, so the error-bound holds.

In state-space, we can interpret this as setting $\dot{x}_2(t) = 0$, and solving for $x_2(t)$.

Example: Balanced singular perturbation

$$\hat{G}(s) = \frac{(s + 10)(s - 5)(s^2 + 2s + 5)(s^2 - 0.5s + 5)}{(s + 4)(s^2 + 4s + 8)(s^2 + 0.2s + 100)(s^2 + 5s + 2000)},$$

transfer function magnitude versus frequency



Optimal Hankel-norm approximation

Given Γ_G , find an operator $\Gamma_{G_r} : L_2(-\infty, 0] \rightarrow L_2[0, -\infty)$ which solves

$$\begin{aligned} & \text{minimize} && \|\Gamma_G - \Gamma_{G_r}\| \\ & \text{subject to} && \Gamma_{G_r} \text{ is the Hankel operator for some } G_r \in H_\infty \\ & && \text{rank}(\Gamma_{G_r}) = r \end{aligned}$$

Notes

- Computational approach similar to balanced truncation.
- The optimal approximant achieves $\|\Gamma_G - \Gamma_{G_r}\| = \sigma_{r+1}$.
- The Hankel-norm is independent of D . If $\sigma_r > \sigma_{r+1}$, one can choose D such that

$$\|G - G_r\| \leq \sigma_{r+1} + \dots + \sigma_n \quad \text{excluding multiplicities}$$

This is half that achieved by balanced truncation.

- For model reduction by 1-state, this is optimal; $\|G - G_r\| \leq \sigma_n$.

Generalized gramians

If $X, Y \in \mathbb{R}^{n \times n}$ and $X = X^T$ and $Y = Y^T$ satisfy the LMIs

$$AX + XA^* + BB^* \leq 0$$

$$A^*Y + YA + C^*C \leq 0$$

then X and Y are called *generalized gramians*.

Notes

- $X \geq X_c$ and $Y \geq Y_o$.
- Generalized gramians can be balanced; in this case the diagonal entries γ_k are called *generalized Hankel singular values*.
- $\gamma_i \geq \sigma_i$ for $i = 1, \dots, n$.
- The generalized gramians can be used for balancing. If $\gamma_r > \gamma_{r+1}$, then the reduced-order model is stable and satisfies

$$\|G - G_r\| \leq 2(\gamma_{r+1} + \dots + \gamma_n) \quad \text{excluding multiplicities}$$

- A useful advantage is that one can search for generalized gramians that increase multiplicities.

Optimal induced-norm model reduction

Suppose G has realization (A, B, C, D) and $A \in \mathbb{R}^{n \times n}$ is Hurwitz. Then the following are equivalent.

- (a) There exists G_r with realization (A_r, B_r, C_r, D_r) of order r such that $\|G - G_r\| < \gamma$.
- (b) There exist $X > 0$ and $Y > 0$ satisfying
 - (i) $AX + XA^* + BB^* < 0$,
 - (ii) $A^*Y + YA + C^*C < 0$,
 - (iii) $\lambda_{\min}(XY) = \gamma^2$, with $\text{rank}(XY - \gamma^2 I) \leq r$.

Notes

- Once X and Y are known, construction is simple, via solving an LMI.
- Problem: the set of X and Y satisfying these constraints is not convex. All known algorithms require computational time $T > c_1 e^{c_2 n}$ for some $c_1, c_2 > 0$.
- General rank-constrained LMIs are known to be NP-complete.
- Good heuristics exist; e.g. minimize $\text{Trace}(XY)$.
- Much more is known; frequency weighted, gap metric, unstable systems, etc.