$\mathbf{H}_2$ Optimal Control

Ulf Jönsson

Optimization and Systems Theory

Royal Institute of Technology

Sweden

References

- Essentials of Robust Control, by Zhou and Doyle
- A course in Robust Control Theory, by Dullerud and Paganini
- State-Space Solutions to Standard $\mathbf{H}_2$ and $\mathbf{H}_\infty$ Control Problems (DGKF), by Doyle, Glover, Khargonekar, and Francis.

- The $\mathbf{H}_2$-norm
- Interpretation of the $\mathbf{H}_2$-norm
- The $\mathbf{H}_2$-optimal control problem
  - Interpretation
  - Proof

Rational $\mathbf{H}_2$-spaces

$\mathbf{RL}_2$ The space of strictly proper real rational transfer matrices with no poles on the imaginary axis. The inner product on $\mathbf{RL}_2$ is defined as

$$\langle F, G\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{tr}(F(j\omega)^* G(j\omega))d\omega = \int_{-\infty}^{\infty} \operatorname{tr}(f(t)^T g(t))dt$$

where g and f are the corresponding impulse response functions (inverse Laplace transforms).

$\mathbf{RH}_2$ The subspace of strictly proper stable transfer function, i.e. all poles are in the open left half plane. The norm is

$$\|G\|_{\mathbf{H}_2} = \frac{1}{2\pi}\sqrt{\int_{-\infty}^{\infty} \operatorname{tr}(G(j\omega)^* G(j\omega))d\omega}$$

$\mathbf{RH}_2^\perp$ The subspace of strictly proper unstable transfer functions, i.e. all poles are in the open right half plane

Lemma 1. *The spaces $\boldsymbol{RH}_2^\perp$ and $\boldsymbol{RH}_2$ are orthogonal.*

Proof. If $G \in \mathbf{RH}_2$ and $F \in \mathbf{RH}_2^\perp$ then

$$\langle G, F\rangle = \int_{-\infty}^{\infty} \operatorname{tr}(g(t)^T f(t))dt = 0$$

since $g(t) = 0$ when t<0 and $f(t) = 0$ when $t > 0$. □

The $\mathbf{H}_2$-norm

Lemma 2. *Let* $G = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *then the $\boldsymbol{H}_2$ norm can be computed as*

$$\|G\|^2_{\boldsymbol{H}_2} = \operatorname{tr}(CPC^T) = \operatorname{tr}(B^TQB)$$

where P and Q are the observability and controllability Grammians defined by the Lyapunov equations

$$AP + PA^T + BB^T = 0$$
$$A^TQ + QA + C^TC = 0$$

Remark 1. Note that the direct term $D = 0$ otherwise $\|G\|_{\mathbf{H}_2} = \infty$.

Proof. We have $g(t) = Ce^{At}B$ for $t \geq 0$ and zero otherwise. Hence,

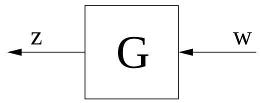
$$\|G\|_{\mathbf{H}_2} = \int_0^\infty \operatorname{tr}(B^Te^{A^Tt}C^TCe^{At}dtB) = \operatorname{tr}(B^TQB)$$

since $Q = \int_0^\infty e^{A^Tt}C^TCe^{At}dt$. To see the last statement note that

$$A^TQ + QA = \int_0^\infty \frac{d}{dt}\left\{e^{A^Tt}C^TCe^{At}\right\}dt = \left[e^{A^Tt}C^TCe^{At}\right]_0^\infty = -C^TC$$

The other part is proven similarly. □

Interpretation of the $\mathbf{H}_2$-norm



Case 1: If w is a Brownian motion (white noise) then

$$dX(t) = AX(t)dt + BdW(t)$$
$$Z(t) = CX(t)$$

where $E\{dW(t)\} = 0$ and $E\{dW(t)dW(s)^T\} = \delta(t-s)Idt$. We get

$\{Z(T\} = 0$ and

$$
\begin{aligned}
E\{|Z(T)|^2\} &= E\{\mathrm{tr}(CX(T)X(T)^TC^T)\}\\
&= \mathrm{tr}(C\int_0^T\int_0^T e^{At}BE\{dW(t)dW(s)^T\}B^Te^{A^Ts}dsdtC^T)\\
&= \mathrm{tr}(C\int_0^T e^{At}BB^Te^{A^Tt}dt \to \mathrm{tr}(CPC^T) = \|G\|^2_{\mathbf{H}_2}
\end{aligned}
$$

as $T \to \infty$, i.e the $\mathbf{H}_2$-norm is the stationary variance.

Case 2: Response to dirac input. If $w_i(t) = \delta(t)e_i$ then if $z_i = Gw_i$ we have

$$\sum_{i=1}^m \|z_i\|_2^2 = \|G\|^2_{\mathbf{H}_2}$$

Proof: Homework.

Case 3: The $\mathbf{L}_2 \to \mathbf{L}_\infty$ induced norm for MISO systems, i.e.,

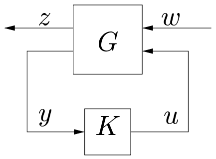
$$\|G\|_{\mathbf{L}_2\to\mathbf{L}_\infty} = \sup_{\|w\|_2\le 1}\|z\|_\infty = \|G\|_{\mathbf{H}_2}$$

where $G \in \mathbf{RH}_2^{1\times m}$. **Proof:**

$$
\begin{aligned}
|z(t)| &= |\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{z}(j\omega)e^{j\omega}d\omega| = |\frac{1}{2\pi}\int_{-\infty}^{\infty}G(j\omega)\hat{w}(j\omega)e^{j\omega}d\omega|\\
&\le \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty}G(j\omega)G(j\omega)^*d\omega}\sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{w}(j\omega)^*\hat{w}(j\omega)d\omega}\\
&= \|G\|_{\mathbf{H}_2}\|w\|_2
\end{aligned}
$$

and we have equality if $\hat{w}(j\omega) = \alpha G(j\omega)^*e^{-j\omega t}$.

The H_2 Optimal Control Problem



Given a state space system $G = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right]$, find a proper rational controller such that

(i) G is internally stabilized

(ii) $\|\mathcal{S}_l(G,K)\|_{\mathbf{H_2}}$ is minimized

Remark 2. $D_{11} = 0$ to ensure that $\mathcal{S}_l(G,K)$ is strictly proper

(otherwise the $\mathbf{H}_2$-norm is infinity.
$D_{22} = 0$ only for convenience.

Theorem 1. *Suppose*

(i) (A, B_1, C_1) *is stabilizable and detectable*

(ii) (A, B_2, C_2) *is stabilizable and detectable*

(iii) $R_1 = D_{12}^T D_{12} > 0$ *and* $R_2 = D_{21} D_{21}^T > 0$

Let $X_2 \geq 0$ *and* $Y_2 \geq 0$ *solve*

$$\begin{aligned} A^T X_2 + X_2 A + C_1^T C_1 &= (X_2 B_2 + C_1^T D_{12}) R_1^{-1} (X_2 B_2 + C_1^T D_{12})^T \\ A Y_2 + Y_2 A^T + B_1 B_1^T &= (Y C_2^T + D_{21} B_1^T) R_2^{-1} (Y C_2^T + D_{21} B_1^T)^T \end{aligned}$$

and let $F_2 = -R_1^{-1}(B_2^T X_2 + D_{12}^T C_1)$, $L_2 = -(Y_2 C_2^T + B_1 D_{21}^T) R_2^{-1}$. *The* $\boldsymbol{H}_2$*-optimal controller is given by*

$$K = \left[\begin{array}{c|c} A + B_2 F_2 + L_2 C_2 & -L_2 \\ \hline F_2 & 0 \end{array}\right]$$

Remark 3. The controller is observer based as was the nominal controller in the Youla parametrization. The state feedback is the solution to the full state LQ problem

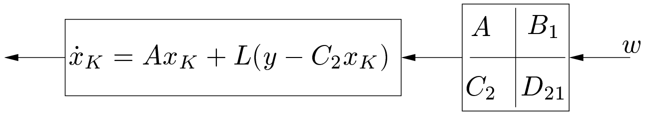
$$\min \int_0^\infty |C_1 x + D_{12} u|^2 dt \quad \text{subj. to} \quad \dot{x} = Ax + B_2 u$$

This means that $Q = C_1^T C_1$, $S = C_1^T D_{12}$, $R = R_1 = D_{12}^T D_{12}$. The solution is given by the state feedback
$u = F_2 x = -R_1^{-1}(B_2^T X_2 + D_{12}^T C_1)x$.

This LQ problem is in turn an $\mathbf{H}_2$ optimal control problem for

$$G = \left[\begin{array}{c|c} A & B_2 \\ \hline C_1 & D_{12} \\ I & 0 \end{array}\right]$$

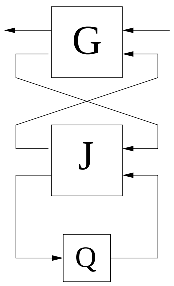
The observer is the optimal Kalman filter for the following system



By duality, the optimal observer gain is obtained as the solution to the $\mathbf{H}_2$ optimal control problem with

$$G = \left[\begin{array}{c|c} A^T & C_2^T \\ \hline B_1^T & D_{21}^T \\ I & 0 \end{array}\right]$$

Proof: The proof is inspired by the Youla parametrization. We will show that $Q = 0$ gives optimal $\mathbf{H}_2$ norm for the system



where $J = \left[\begin{array}{c|cc} A + B_2 F_2 + L_2 C_2 & -L_2 & B_2 \\ \hline F_2 & 0 & I \\ -C_2 & I & 0 \end{array}\right]$

e have the closed loop transfer function

$$T_{cl} = \mathcal{S}_l(T, Q) = T_{11} + T_{12} Q T_{21}$$

where

$$T = G \star J = \left[\begin{array}{cc|cc} A + B_2F_2 & -B_2F_2 & B_1 & B_2 \\ 0 & A + L_2C_2 & B_1 + L_2D_{21} & 0 \\ \hline C_1 + D_{12}F_2 & -D_{12}F_2 & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right]$$

Hence

$$T_{12} = \left[\begin{array}{c|c} A + B_2F_2 & B_2 \\ \hline C_1 + D_{12}F_2 & D_{12} \end{array}\right] \quad T_{21} = \left[\begin{array}{c|c} A + L_2C_2 & B_1 + L_2D_{21} \\ \hline C_2 & D_{21} \end{array}\right]$$

$$T_{11} = \left[\begin{array}{cc|c} A + B_2F_2 & -B_2F_2 & B_1 \\ 0 & A + L_2C_2 & B_1 + L_2D_{21} \\ \hline C_1 + D_{12}F_2 & -D_{12}F_2 & 0 \end{array}\right]$$

$$= \left[\begin{array}{c|c} A + B_2F_2 & B_1 \\ \hline C_1 + D_{12}F_2 & 0 \end{array}\right] - \left[\begin{array}{c|c} A + B_2F_2 & B_2F_2 \\ \hline C_1 + D_{12}F_2 & D_{12}F_2 \end{array}\right]$$

$$\times \left[\begin{array}{c|c} A + C_2L_2 & B_2L_2 \\ \hline I & 0 \end{array}\right]$$

here is a point in further factorizing these transfer functions. Let

$$G_c = \left[\begin{array}{c|c} A + B_2F_2 & I \\ \hline C_1 + D_{12}F_2 & 0 \end{array}\right], \qquad G_f = \left[\begin{array}{c|c} A + C_2L_2 & B_2L_2 \\ \hline I & 0 \end{array}\right]$$

$$U = \left[\begin{array}{c|c} A + B_2F_2 & B_2R_1^{-1/2} \\ \hline C_1 + D_{12}F_2 & D_{12}R_1^{-1/2} \end{array}\right], \quad V = \left[\begin{array}{c|c} A + L_2C_2 & B_1 + L_2D_{21} \\ \hline R_2^{-1/2}C_2 & R_2^{-1/2}D_{21} \end{array}\right]$$

Hence

$$T_{11} = G_cB_1 - UR_1^{1/2}F_2G_f$$
$$T_{12} = UR_1^{1/2}$$
$$T_{21} = R_2^{1/2}V$$

mma 3. *We have*

(i) U *is inner, i.e.* $U^*U = I$

(ii) V *is co-inner, i.e.* $VV^* = I$

(iii) $U^*G_c \in \boldsymbol{RH}_2^{\perp}$

(iv) $G_fV^* \in \boldsymbol{RH}_2^{\perp}$

Proof. Let $A_{F2} = A + B_2F_2$ and $C_{F2} = C_1 + D_{12}F_2$. Since the adjoint is $U^* = \left[\begin{array}{c|c} -A_{F_2}^T & -C_{F_2}^T \\ \hline R_1^{-1/2}B_2^T & R_1^{-1/2}D_{12}^T \end{array}\right]$ we get

$$U^*U = \left[\begin{array}{cc|c} -A_{F_2}^T & -C_{F_2}^TC_{F_2} & -C_{F_2}^TD_{12}R_1^{-1/2} \\ 0 & A_{F_2} & B_2R_1^{-1/2} \\ \hline R_1^{-1/2}B_2^T & R_1^{-1/2}D_{12}^TC_{F_2} & I \end{array}\right]$$

d

$$U^*G_c = \left[\begin{array}{cc|c} -A_{F_2}^T & -C_{F_2}^T C_{F_2} & 0 \\ 0 & A_{F_2} & I \\ \hline R_1^{-1/2}B_2^T & R_1^{-1/2}D_{12}^T C_{1F_2} & 0 \end{array}\right]$$

Next we make a similarity transformation using $M = \begin{bmatrix} I & -X_2 \\ 0 & I \end{bmatrix}$.

We get

$$U^*U = \left[\begin{array}{c|c} M\bar{A}M^{-1} & M\bar{B} \\ \hline \bar{C}M^{-1} & \bar{D} \end{array}\right] = \left[\begin{array}{cc|c} -A_{F_2}^T & Q_1 & Q_2 \\ 0 & A_{F_2} & B_2R_1^{-1/2} \\ \hline R_1^{-1/2}B_2^T & Q_2^T & I \end{array}\right]$$

where $Q_1 = -(A_{F_2}^T X_2 + X_2A_{F_2} + C_{F_2}^T C_{F_2})$ and
$Q_2 = -(C_{F_2}^T D_{12} + X_2B_2)R_1^{-1/2}$.

ote that from the Riccati eqn and the definition of F_2 we have

$$A^TX_2 + X_2A + C_1C_1^T = (X_2B_2 + C_1^TD_{12})R_1^{-1}(X_2B_2 + C_1^TD_{12})^T$$

$$\Longleftrightarrow$$

$$A_{F_2}^TX_2 + X_2A_{F_2} + C_{F_2}^TC_{F_2} = 0$$

$$\Longleftrightarrow$$

$$Q_1 = 0$$

and

$$F_2 = -R_1^{-1}(B_2^TX_2 + D_{12}^TC_1 \quad \Longleftrightarrow \quad Q_2 = 0$$

Hence

$$U^*U = \left[\begin{array}{cc|c} -A_{F_2}^T & 0 & 0 \\ 0 & A_{F_2} & B_2R_1^{-1/2} \\ \hline R_1^{-1/2}B_2^T & 0 & I \end{array}\right] = I$$

imilarly,

$$U^*G_c = \left[\begin{array}{cc|c} -A_{F_2}^T & 0 & -X_2 \\ 0 & A_{F_2} & I \\ \hline R_1^{-1/2}B_2^T & 0 & I \end{array}\right] = \left[\begin{array}{c|c} -A_{F_2}^T & -X_2 \\ \hline R_1^{-1/2}B_2^T & 0 \end{array}\right] \in \mathbf{RH}_2^\perp$$

A dual argument shows $VV^* = I$ and $G_fV^* \in \mathbf{RH}_2^\perp$. By using

- $\langle Uv, u\rangle = \langle v, U^*u\rangle$ (U^* is called the adjoint operator)
- $\|Uv\|_{\mathbf{H}_2}^2 = \langle Uv, Uv\rangle = \langle v, U^*Uv\rangle = \langle v, v\rangle = \|v\|_{\mathbf{H}_2}^2$
- $\langle F, G\rangle = 0$ when $G \in \mathbf{RH}_2$ and $F \in \mathbf{RH}_2^\perp$
- $\|G_1 + G_2\|_{\mathbf{H}_2}^2 = \|G_1\|_{\mathbf{H}_2}^2 + 2\langle G_1, G_2\rangle + \|G_2\|_{\mathbf{H}_2}^2$
- $\langle G_c, UG\rangle = \langle U^*G_c, G\rangle = 0$ since $U^*G_c \in \mathbf{RH}_2^\perp$
- $\langle GV, G_f\rangle = \langle G, G_fV^*\rangle = 0$ since $G_fV^* \in \mathbf{RH}_2^\perp$

.c then we obtain

$$\begin{aligned}
\|T_{cl}\|_{\mathbf{H}_2}^2 &= \|G_cB_1 - UR_1^{1/2}F_2G_f + UR_1^{1/2}QR_2^{1/2}V\|_{\mathbf{H}_2}^2 \\
&= \|G_cB_1\|_{\mathbf{H}_2}^2 + 2\left\langle U^*G_cB_1, (R_1^{1/2}QR_2^{1/2}V - R_1^{1/2}F_2G_f\right\rangle \\
&\quad + \|U(R_1^{1/2}QR_2^{1/2}V - R_1^{1/2}F_2G_f)\|_{\mathbf{H}_2}^2 \\
&= \|G_cB_1\|_{\mathbf{H}_2}^2 + \|(R_1^{1/2}(QR_2^{1/2}V - F_2G_f)\|_{\mathbf{H}_2}^2 \\
&= \|G_cB_1\|_{\mathbf{H}_2}^2 + \|R_1^{1/2}F_2G_f\|_{\mathbf{H}_2}^2 \\
&\quad - \left\langle F_2^TR_1QR_2^{1/2}, G_fV^*\right\rangle + \|R_1^{1/2}QR_2^{1/2}V\|_{\mathbf{H}_2}^2 \\
&= \|G_cB_1\|_{\mathbf{H}_2}^2 + \|R_1^{1/2}F_2G_f\|_{\mathbf{H}_2}^2 + \|R_1^{1/2}QR_2^{1/2}V\|_{\mathbf{H}_2}^2
\end{aligned}$$

The minimum $\mathbf{H}_2$-norm is obtained when $Q = 0$. This proves the theorem. $\square$