GIAN course: Singular optimal control

Slide collection 7 (unfinished)

Schur complement,
Linear quadratic minimization
and
linear quadratic optimal control

Michael Karow, TU Berlin, December 2016

karow@math.tu-berlin.de

The Schur complement

Let $A \in \mathbb{C}^{n\times n}$ be nonsingular. Then, for any Matrix M with upper left block A:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix} \qquad \text{(block } LU\text{-decomposition)}$$

$$= \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}$$

The matrix $D - CA^{-1}B =: M/A$ is called the Schur complement of A in M.

Analogously we have for nonsingular D and the Schur complement $M/D := A - BD^{-1}C$:

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} M/D & 0 \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} M/D & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I & 0 \\ D^{-1}C & I \end{bmatrix}.$$

The Schur complement II

From the identity

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & M/A \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}, \qquad M/A = D - CA^{-1}B,$$

and

$$\det\left(\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}\right) = \det\left(\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}\right) = 1, \quad \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}, \quad \begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -Y \\ 0 & I \end{bmatrix}$$

it follows that

- $\det(M) = \det(A)\det(M/A)$
- $\text{rank}(M) = (\text{size of } A) + \text{rank}(M/A)$
- M nonsingular $\Leftrightarrow$ M/A nonsingular
- $$M^{-1} = \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & (M/A)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix}$$
$$= \begin{bmatrix} A^{-1} + A^{-1}B(M/A)^{-1}CA^{-1} & -A^{-1}B(M/A)^{-1} \\ -(M/A)^{-1}CA^{-1} & (M/A)^{-1} \end{bmatrix}$$

The Schur complement III

Transfer function as Schur complement.

Let $G(s) = D + C(s\,E - A)^{-1}B$, $s \in \mathbb{C}$, be the transfer function of the system

$$E\,\dot{x} = A\,x + B\,u, \quad y = C\,x + B\,u.$$

Then if s is not an eigenvalue of (E, A),

$$\begin{bmatrix} A - s\,E & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ -C(s\,E - A)^{-1} & I \end{bmatrix} \begin{bmatrix} A - s\,E & 0 \\ 0 & G(s) \end{bmatrix} \begin{bmatrix} I & -(s\,E - A)^{-1}B \\ 0 & I \end{bmatrix}.$$

The Schur complement VI

Schur complement and congruence transformation. Let

$$M = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix}, \qquad A = A^*,\ D = D^*,\ A \text{ nonsingular}$$

be Hermitian. Then

$$\begin{aligned} M &= \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ B^*A^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & M/A \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}, \qquad M/A = D - B^*A^{-1}B, \\ &= \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}^* \begin{bmatrix} A & 0 \\ 0 & M/A \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix} \qquad \text{(Congruence transformation)} \end{aligned}$$

Consequences:

(1) $M > 0$ iff ($A > 0$ and $M/A > 0$).

(2)

$$\begin{bmatrix} x \\ y \end{bmatrix}^* \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = (x + A^{-1}B\,y)^* A (x + A^{-1}B\,y) + y^*(M/A)y \qquad \text{(sum of 2 squares)}$$

Linear Quadratic Minimization

Minimization of a linear-quadratic function

Setting: We consider the linear-quadratic functional

$$J : \mathcal{Y} \to \mathbb{R}, \qquad J(y) := \alpha(y,y) - 2\beta(y) + \gamma,$$

where $\mathcal{Y}$ is a vector space over $\mathbb{R}$, $\alpha : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$ is a symmetric bilinear form and $\beta : \mathcal{Y} \to \mathbb{R}$ is a linear form. Furthermore, let $\mathcal{Z}$ be vector space over $\mathbb{R}$, let $G : \mathcal{Y} \to \mathcal{Z}$ be a linear operator, and let $z \in \mathcal{Z}$.

Unconstrained minimization problem (UMP):

Find $y_0 \in \mathcal{Y}$ such that $J(y_0) = \min\{J(y)|\ y \in \mathcal{Y}\}$.

Constrained minimization problem (CMP):

Find $y_0 \in \mathcal{Y}$ such that $Gy_0 = z$ and $J(y_0) = \min\{J(y)|\ y \in \mathcal{Y},\ Gy = z\}$.

Remark: we will not prove nonconstructive existence statements about y_0.
$\Rightarrow$ no topology (completeness of spaces , continuity of functions) needed.

Unconstraint minimization: General facts I

(UMP): minimize $J(y) = \alpha(y,y) - 2\,\beta(y) + \gamma, \quad y \in \mathcal{Y}$.

The function J along a line (Taylor-expansion): For $y_0, \eta \in \mathcal{Y}$, $t \in \mathbb{R}$,

$$\begin{aligned} J(y_0 + t\eta) &= \alpha(y_0 + t\eta, y_0 + t\eta) - 2\beta(y_0 + t\eta)) + \gamma \\ &= \alpha(y_0, y_0) + 2t\alpha(\eta, y_0) + t^2\alpha(\eta,\eta) - 2\beta(y_0) - 2t\beta(\eta) + \gamma \\ &= J(y_0) + 2\,(\alpha(\eta, y_0) - \beta(\eta))\, t + \alpha(\eta,\eta)\, t^2. \qquad \text{(quadratic polynomial in } t\text{)} \end{aligned}$$

Thus:

- $\frac{d}{dt}\Big|_{t=0} J(y_0 + t\eta) = 2\,(\alpha(\eta, y_0) - \beta(\eta))$. (derivative of J in direction η)
- If $\alpha(\eta,\eta) \neq 0$ then
$$J(y_0 + t\,\eta) = \alpha(\eta,\eta)\left(t - \frac{\alpha(\eta, y_0) - \beta(\eta)}{\alpha(\eta,\eta)}\right)^2 + J(x_0) - \frac{(\alpha(\eta, y_0) - \beta(\eta))^2}{\alpha(\eta,\eta)}.$$
- If $\alpha(\eta,\eta) > 0$ the minimum on the line is uniquely attained at $t = \frac{\alpha(\eta,y_0) - \beta(\eta)}{\alpha(\eta,\eta)}$.
- If $\alpha(\eta,\eta) < 0$ then $\lim_{t \to \pm\infty} J(y_0 + t\,\eta) = -\infty$.

(UM)-Proposition: If $\alpha(\eta,\eta) < 0$ for some $\eta \in \mathcal{Y}$, then $\inf_{y \in \mathcal{Y}} J(y) = -\infty$.
Suppose that α is positive semidefinite. Then $y_0 \in \mathcal{Y}$ is a global minimzer for J iff

$$\alpha(\eta, y_0) = \beta(\eta) \quad \text{for all} \quad \eta \in \mathcal{Y}.$$

The minimum then is $J(y_0) = \gamma - \beta(y_0)$.
If α is positive definite then there is at most one minimizer.

Unconstraint minimization: General facts II

Additonal remarks:

- In the semidefinite case a necessary condition for $\inf_{y\in\mathcal{Y}} J(y) \neq -\infty$ is that
$$\beta(\eta) = 0 \quad \text{if} \quad \alpha(\eta,\eta) = 0.$$
Proof: homework.

- If $\mathcal{Y}$ is endowed with a norm $\|\cdot\|$ and there are constants $c, d > 0$ such that for all η,
$$\alpha(\eta,\eta) \geq c\,\|y\|^2, \quad \beta(\eta) \leq d\,\|y\|, \qquad (*)$$
then
$$J(y) = \alpha(y,y) - 2\,\beta(y) + \gamma \geq c\|y\|^2 - 2d\|y\| + \gamma = c(\|y\| - (d/c))^2 + \gamma - d^2/c$$
is bounded below.

- If $\mathcal{Y}$ is a Hilbert space and $(*)$ holds then a unique minimizer y_0 for J exists (Lax-Milgram Theorem).
Proof: see e.g. books on the Finite Element Method.

Unconstraint minimization: the matrix case I

Problem: minimize the linear-quadratic function

$$J(x) = x^*Ax - 2\,\Re(b^*x) + c, \quad A = A^* \in \mathbb{C}^{n\times n}, \quad b \in \mathbb{C}^n, \quad c \in \mathbb{R}.$$

Observation: $x^*Ax = \alpha(x,x)$, where $\alpha(x_1,x_2) = \Re(x_1^*Ax_2)$ (real bilinear)

Corollary to (UM)-Proposition:

- If A is not positive semidefinite then $\inf_{x\in\mathbb{C}^n} J(x) = -\infty$.
- If A is pos. semidef. then $x_0 \in \mathbb{C}^n$ is a minimizer iff $\Re(\xi^*Ax_0) = \Re(\xi^*b)$ for all ξ, i.e.

  $$Ax_0 = b.$$

  The minimum then is $J(x_0) = c - \Re(b^*x_0)$.
- If A is positive definite then the unique minimizer of J is

  $$x_0 = A^{-1}b.$$

  with associated minimum $J(x_0) = c - b^*A^{-1}b$.

Remark: Result shows that in the semidefinite case the problem of solving the equation $Ax_0 = b$ is equivalent to the minimization problem above. This fact is used in Numerical Analysis (CG-Algorithm).

Unconstraint minimization: the matrix case II

Problem: minimize the linear-quadratic function

$$J(x) = x^* A x - 2\,\Re(b^* x) + c, \quad A = A^* \in \mathbb{C}^{n\times n}, \quad b \in \mathbb{C}^n, \quad c \in \mathbb{R}.$$

Solution by completing the square: Suppose A is nonsingular. Then

$$\begin{aligned} J(x) &= \begin{bmatrix} x \\ 1 \end{bmatrix}^* \begin{bmatrix} A & -b \\ -b^* & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ 1 \end{bmatrix}^* \overbrace{\begin{bmatrix} I & -A^{-1}b \\ 0 & I \end{bmatrix}^* \begin{bmatrix} A & 0 \\ 0 & c - b^*A^{-1}b \end{bmatrix} \begin{bmatrix} I & -A^{-1}b \\ 0 & I \end{bmatrix}}^{\text{Schur complement factorization}} \begin{bmatrix} x \\ 1 \end{bmatrix} \\ &= \underbrace{(x - A^{-1}b)^* A (x - A^{-1}b)}_{\text{quadratic term}} + \underbrace{c - b^*A^{-1}b}_{\text{constant}}. \end{aligned}$$

Conclusion:

- If A is positive definite then $\min(J) = c - b^*A^{-1}b$, $\quad$ $\text{argmin}(J) = A^{-1}b$
- If A is not positive definite then J has no minimum.

  Reason: Let $x_t = A^{-1}b + t\,\xi$, where $\xi^* A \xi < 0$ and $t \in \mathbb{R}$. Then

$$J(x_t) = t^2\,\xi^* A \xi + c - b^* A^{-1} b \to -\infty \quad \text{as} \quad t \to \infty.$$

Unconstraint minimization: least squares

Problem: minimize

$$J(x) = \|Ax - b\|_2^2, \quad A \in \mathbb{C}^{m\times n}, \text{ (rectangular)} \quad b \in \mathbb{C}^m.$$

Observation:

$$\begin{aligned} J(x) &= (Ax-b)^*(Ax-b) = (Ax)^*(Ax) - (Ax)^*b - b^*(Ax) + b^*b \\ &= \underbrace{x^*A^*Ax}_{\geq 0} - 2\,\Re(x^*A^*b) + \|b\|_2^2. \end{aligned}$$

Corollary to (UM)-Proposition:

$$\begin{aligned} x_0 \text{ minimizes } J \quad &\Leftrightarrow \quad \forall \xi:\ \xi^*(A^*A)x_0 = \xi^*A^*b \quad \Leftrightarrow \quad \forall \xi:\ (A\xi)^*(Ax_0 - b) = 0 \qquad (*) \\ &\Leftrightarrow \quad A^*Ax_0 = A^*b. \qquad \text{(normal equation)} \end{aligned}$$

Minimum: $J(x_0) \;=\; \|Ax_0 - b\|_2^2 = \|b\|_2^2 - \Re(x_0^*A^*b).$

Interpretation: $\sqrt{J(x_0)}$ is the Euclidean distance of the vector b from the vector space $\operatorname{Im} A$, i.e. the space spanned by the columns of A. Ax_0 is the point in that space which is the nearest to b. By $(*)$ the vector $Ax_0 - b$ is orthogonal to all vectors in $\operatorname{Im} A$. Thus: Ax_0 is the orthonal projection of b onto $\operatorname{Im} A$.

The constraint minimization problem: general facts

(CMP): minimize $J(y) = \alpha(y,y) - 2\beta(y) + \gamma$ subject to $Gy = z$.

(CM)-Proposition: (Notation: $\ker G = \{y \in \mathcal{Y} | \, G(y) = 0\}$).

(1) If there exists $\eta \in \ker G$ with $\alpha(\eta,\eta) < 0$ then (CMP) has no solution.

(2) Suppose $\alpha(\eta,\eta) \geq 0$ for all $\eta \in \ker G$. Then y_0 is a minimizer subject to $Gy = z$ iff

$$Gy_0 = z, \qquad \text{and} \qquad \alpha(\eta, y_0) = \beta(\eta) \quad \text{for all } \eta \in \ker G. \qquad (*)$$

(3) Suppose $\alpha(\eta,\eta) \geq 0$ for all $\eta \in \ker G$. If $y_0 \in \mathcal{Y}$ and the linear form ℓ satisfy

$$Gy_0 = z, \qquad \text{and} \qquad \alpha(\eta, y_0) + \ell(G\eta) = \beta(\eta) \quad \text{for all } \eta \in \mathcal{Y} \qquad (**)$$

then y_0 is a minimizer for (CM). The minimum is

$$J(y_0) = \gamma - \ell(z) - \beta(y_0). \qquad (***)$$

Proof: Let $Gy_0 = z$, $\eta \in \ker G$, $y_t = y_0 + t\eta$, $t \in \mathbb{R}$. Then $Gy_t = z$. We have,

$$J(y_t) = J(y_0) + 2(\alpha(\eta, y_0) - \beta(\eta))\, t + \alpha(\eta,\eta)\, t^2.$$

This yields (1) and (2). $(*)$ implies $(**)$. Replacing η by y_0 in $(**)$ yields $(***)$. □

Remark: The linear form ℓ is called Lagrangian multiplier. Idea behind (3):
For any linear ℓ, the functional J and its modification

$$J^\ell(y) := J(y) + 2\,\ell(Gy - z), \qquad y \in \mathcal{Y}$$

have the same values if $Gy = z$.
Searching for (y_0, ℓ) such that $(**)$ holds is often easier than searching for y_0 directly.

The constraint minimization problem: general facts II

(CMP): minimize $J(y) = \alpha(y,y) - 2\beta(y) + \gamma$ subject to $Gy = z$.

Supplement to (CM)-Proposition:

Suppose $\alpha(\eta,\eta) \geq 0$ for all $\eta \in \ker G$.
On the foregoing slide we discussed the conditions

(∗) $Gy_0 = z$, and $\alpha(\eta, y_0) = \beta(\eta)$ for all $\eta \in \ker G$.

(∗∗) $Gy_0 = z$, and $\alpha(\eta, y_0) + \ell(G\eta) = \beta(\eta)$ for all $\eta \in \mathcal{Y}$

Condition (∗) is necessary and sufficient for a minimizer. (∗∗) implies (∗).

If $\mathcal{Y}, \mathcal{Z}$ are Hilbert spaces endowed with respect to inner products $\langle \cdot, \cdot \rangle$ and α, β, G are continuous then (∗) implies (∗∗). More precisely, for some $\zeta \in \mathcal{Z}$,

$$\ell(G\eta) = \langle \eta, G^*\zeta \rangle,$$

where G^* is the adjoint of G.

Proof sketch. To the continuous linear functional $d(\eta) := \beta(\eta) - \alpha(\eta, y_0)$ there exists w such that $d(\eta) = \langle \eta, w \rangle$. By (∗), $w \in (\ker G)^\perp$.
In Hilbert space we have for continuous G, $(\ker G)^\perp = \operatorname{Im} G^*$. Thus, $w = G^*\zeta$.

Constraint minimization: the matrix case I

Problem: minimize

$$J(x) = x^* A x - 2\,\Re(x^* b) + c, \qquad A = A^* \in \mathbb{C}^{n\times n}, \quad b \in \mathbb{C}^n, \quad c \in \mathbb{R}.$$

subject to

$$Gx = z, \qquad G \in \mathbb{C}^{m\times n}, \;\; z \in \mathbb{C}^m.$$

Corollary to (CM)-Proposition and its supplement:

- If A is not positive semidefinite on $\ker G$ then there is no minimum.
- Suppose A is positive semidefinite on $\ker G$.

  Then the following statements are equivalent.

  (1) x_0 is a minimizer.

  (2) $Gx_0 = z$, and for all $\xi \in \ker G$: $\Re(\xi^* A x_0) = \Re(\xi^* b)$.

  (3) $Gx_0 = z$, and there is a $\zeta \in \mathbb{C}^m$ such that for all $\xi \in \mathbb{C}^n$

  $$\Re(\xi^* A x_0) + \Re(\xi^* G^* \zeta) = \Re(\xi^* b).$$

  (4) There is a $\zeta \in \mathbb{C}^m$ such that

  $$\begin{bmatrix} 0 & G \\ G^* & A \end{bmatrix} \begin{bmatrix} \zeta \\ x_0 \end{bmatrix} = \begin{bmatrix} z \\ b \end{bmatrix}.$$

  If these conditions are satisfied then the minimum is $J(x_0) = c - \Re(z^* \zeta) - \Re(x_0^* b)$.

Constraint minimization: the matrix case II

Problem: minimize

$$J(x) = x^*Ax - 2\,\Re(x^*b) + c, \qquad A = A^* \in \mathbb{C}^{n\times n}, \quad b \in \mathbb{C}^n, \quad c \in \mathbb{R}.$$

subject to

$$Gx = z, \qquad G \in \mathbb{C}^{m\times n}, \;\; z \in \mathbb{C}^m.$$

Solution by completing the square.

Suppose $\begin{bmatrix} 0 & G \\ G^* & A \end{bmatrix}$ is nonsingular. Then for all x with $Gx = z$ and all ζ,

$$\begin{aligned} J(x) &= J(x) + 2\Re(\zeta^*(Gx - z)) = \begin{bmatrix} \zeta \\ x \end{bmatrix}^* \begin{bmatrix} 0 & G \\ G^* & A \end{bmatrix} \begin{bmatrix} \zeta \\ x \end{bmatrix} - 2\,\Re(\begin{bmatrix} \zeta \\ x \end{bmatrix}^* \begin{bmatrix} z \\ b \end{bmatrix}) + c \\ &= \left(\begin{bmatrix} \zeta \\ x \end{bmatrix} - \begin{bmatrix} 0 & G \\ G^* & A \end{bmatrix}^{-1} \begin{bmatrix} z \\ b \end{bmatrix}\right)^* \begin{bmatrix} 0 & G \\ G^* & A \end{bmatrix} \left(\begin{bmatrix} \zeta \\ x \end{bmatrix} - \begin{bmatrix} 0 & G \\ G^* & A \end{bmatrix}^{-1} \begin{bmatrix} z \\ b \end{bmatrix}\right) + c - \begin{bmatrix} z \\ b \end{bmatrix}^* \begin{bmatrix} 0 & G \\ G^* & A \end{bmatrix}^{-1} \begin{bmatrix} z \\ b \end{bmatrix} \end{aligned}$$

Now let

$$\begin{bmatrix} \zeta \\ x_0 \end{bmatrix} := \begin{bmatrix} 0 & G \\ G^* & A \end{bmatrix}^{-1} \begin{bmatrix} z \\ b \end{bmatrix}.$$

Then

$$J(x) = (x - x_0)^* A (x - x_0) + c - \zeta^* z - x_0^* b.$$

If A is positive semidefinite on $\ker G$ then the minimum is attained for $x = x_0$.

Constraint minimization: quadratic dependence of infimum on constraint

We consider a pure quadratic minimization problem with a linear constraint. Let

$$V(z) := \inf\{\alpha(y,y) \mid \ y \in \mathcal{Y}, \ G(y) = z\},$$

where α is symmetric bilinear and $G : \mathcal{Y} \to \mathcal{Z}$ is linear.

Proposition: Let $\mathcal{S} \subseteq \mathcal{Z}$ be a subspace such that

$$-\infty < V(z) \text{ for all } z \in \mathcal{S}.$$

Then there exists a symmetric bilinear form $\pi : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ such that

$$V(z) = \pi(z,z) \text{ for all } z \in \mathcal{S}.$$

If $\mathcal{Y}, \mathcal{S}$ are vectors spaces over $\mathbb{C}$ and α is Hermitian then π is also Hermitian, in particular, if $\mathcal{S} = \mathbb{C}^n$,

$$V(z) = \pi(z,z) = z^* P z,$$

for some Hermitian matrix $P \in \mathbb{C}^{n\times n}$.

The proof is based on the lemma on the following page.

Constraint minimization: quadratic dependence of infimum on constraint

Lemma. Suppose the function $q : \mathcal{S} \to \mathbb{R}$ satisfies for all for all $z, w \in \mathcal{S}, \lambda \in \mathbb{K}$,

(1) $q(z+w, z+w) + q(z-w, z-w) = 2\,(q(z,z) + q(z,w))$,

(parallelogram identity)

(2) $q(z + \lambda\, w) \to q(z)$ as $\lambda \to 0$,

Then **(polarization identities)**

(i) $\pi_0(z,w) := \frac{1}{4}(q(z+w, z+w) - q(z-w, z-w))$ is bilinear.

(ii) If $\mathcal{S}$ is a vector space over $\mathbb{C}$, then

$$\pi(z,w) = \frac{1}{4}(q(z+w, z+w) - q(z-w, z-w) - i\, q(i\, z + w, i\, z + w) + i\, q(i\, z - w, i\, z - w))$$

is Hermitian.

Furthermore, $q(z) = \pi_0(z,z) = \pi(z,z)$.

Proof. Hint: symmetry is trivial. Show additivity. Using additivity show linearity for integers, then for rational scalars, then using (2) for all reals.

Sketch of proof of Proposition. $q(z) := \alpha(z,z)$ satisfies (1),(2). Thus, definition of V implies

$$V(z+w) + V(z-w) \leq 2(V(z) + V(w) + 2\epsilon), \qquad \epsilon > 0.$$

Thus $V(z+w) + V(z-w) \leq 2(V(z) + V(w))$.

Replacing z by $z+w$ and w by $z-w$ yields (1). proof of (2) for V is analogous.

Linear Quadratic Optimal Control

Finite Time Horizon

Author's advice:

Read the old masters

Jan C. Willems:

Least Squares Stationary Optimal Control and the Algebraic Riccati Equation.

IEEE Transactions on Automatic Control 16, Nr. 6 (1971)

Roger W. Brockett:

Book: Finite Dimensional Linear Systems. Wiley, 1970.

Bryson, Ho:

Book: Applied optimal control. Taylor & Francis, 1975.

Book: Jacobson, Speyer:

Primer on Optimal Control Theory. SIAM 2010.

The linear quadratic optimal control. Finite time horizon, general case

Linear-quadratic control problem: Minimize

$$J(x,u) = x(T)^* P_{_T}\, x(T) - 2\,\Re(x(T)^* b_{_T}) + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} - 2\,\Re(\begin{bmatrix} q \\ r \end{bmatrix}^* \begin{bmatrix} x \\ u \end{bmatrix})\, dt$$

subject to

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \quad \Psi_{_T}\, x(T) = c. \qquad (*)$$

The final time $T \in (0, \infty)$ is fixed. The control u is a piecewise continuos function of t. The state x is a continuous and piecewise differentiable function of t. All matrices and vectors in the integral as well as A, B may be time dependent and piecewise continuous. The matrices $P_{_T}, Q$ and R are Hermitian. The matrix $\Psi_{_T}$ may be non quadratic (it may be a row vector). If $\Psi_{_T} = 0$, $c = 0$, then the end condition is void.

Observation. J is linear-quadratic: $J(x,u) = \alpha\left(\begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} x \\ u \end{bmatrix}\right) - 2\,\beta\left(\begin{bmatrix} x \\ u \end{bmatrix}\right)$, where

$$\alpha\left(\begin{bmatrix} x_1 \\ u_1 \end{bmatrix}, \begin{bmatrix} x_2 \\ u_2 \end{bmatrix}\right) := x_1(T)^* P_{_T}\, x_2(T) + \int_0^T \begin{bmatrix} x_1 \\ u_1 \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x_2 \\ u_2 \end{bmatrix} dt,$$

$$\beta\left(\begin{bmatrix} x \\ u \end{bmatrix}\right) := \Re\left(b_{_T}^* x(T) + \int_0^T \begin{bmatrix} q \\ r \end{bmatrix}^* \begin{bmatrix} x \\ u \end{bmatrix} dt\right).$$

The constraints $(*)$ can be written as

$$G\left(\begin{bmatrix} x \\ u \end{bmatrix}\right) = \begin{bmatrix} x_0 \\ c \\ 0 \end{bmatrix}, \qquad \text{where} \quad G\left(\begin{bmatrix} x \\ u \end{bmatrix}\right) := \begin{bmatrix} x(0) \\ \Psi_{_T} x(T) \\ Ax + Bu - \dot{x} \end{bmatrix}.$$

Special case: the tracking problem

Cost functionals with linear terms appear in the following **tracking problem:**

Linear system with output y

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \qquad y = Cx.$$

Output should follow the given **reference trajectory** y_r.

Cost functional ($\gamma_1, \gamma_2, \gamma_3 \geq 0$ given constants, $\|\cdot\|$ Euclidean norm):

$$\begin{aligned} J(x,u) \;&:=\; \gamma_1 \|y(T) - y_r(T)\|^2 + \int_0^T \gamma_1 \|y - y_r\|^2 + \gamma_3 \|u\|^2 \, dt \\ &=\; \gamma_1 \left(x(T)^* C(T)^* C(T) x(T) - 2\Re(x(T)^* C(T)^* y_r(T)) + \|y_r(T)\|_2^2\right) \\ &\quad + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} \gamma_2 C^* C & \gamma_2 C \\ \gamma_2 C^* & \gamma_3 I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} - 2\,\Re(\begin{bmatrix} \gamma_2 C^* y_r \\ 0 \end{bmatrix}^* \begin{bmatrix} x \\ u \end{bmatrix}) + \gamma_2 \|y_r\|^2 \, dt \\ &\quad + \int_0^T \gamma_3 \|u\|^2 \, dt \end{aligned}$$

Basic case

Simple control problem (pure quadratic, no end condition): Minimize

$$\hat{J}(x,u) = \alpha\left(\begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} x \\ u \end{bmatrix}\right) = x(T)^* P_{\scriptscriptstyle T} x(T) + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \underbrace{\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix}}_{\mathcal{Q}} \begin{bmatrix} x \\ u \end{bmatrix} dt$$

subject to

$$G\left(\begin{bmatrix} x \\ u \end{bmatrix}\right) = \begin{bmatrix} x(0) \\ Ax + Bu - \dot{x} \end{bmatrix} = \begin{bmatrix} x_0 \\ 0 \end{bmatrix}. \qquad \text{(no end constraint)}$$

According to (CLQM)-Proposition: If

(1) $G\left(\begin{bmatrix} h \\ v \end{bmatrix}\right) = 0$ implies $\alpha\left(\begin{bmatrix} h \\ v \end{bmatrix}, \begin{bmatrix} h \\ v \end{bmatrix}\right) \geq 0$;

(2) there is a linear functional ℓ such that $\alpha\left(\begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} h \\ v \end{bmatrix}\right) + \ell\left(G\left(\begin{bmatrix} h \\ v \end{bmatrix}\right)\right) = 0$

for all piecewise differentiable h and all piecewise continuous v

then

(x,u) is a minimizer, and the minimum is $J(x,u) = -\ell(\begin{bmatrix} x_0 \\ 0 \end{bmatrix})$.

Note: If (1) is not satisfied, then there is no minimum.
If $P_{\scriptscriptstyle T}$ and $\mathcal{Q}$ are positive semidefinite then (1) trivially holds.

Ansatz for linear functional: $\ell\left(\begin{bmatrix} w_0 \\ e \end{bmatrix}\right) = w_0^* \nu_0 + \int_0^T e^* \lambda \, dt,$

where ν_0 is a vector and λ is a continuos and piecewise differentiable function (Lagrange multipliers) which have to be determined.

Basic case: derivation of solution I

We have to satisfy

$$0 = \alpha\left(\begin{bmatrix} h \\ v \end{bmatrix}, \begin{bmatrix} x \\ u \end{bmatrix}\right) + \ell\left(G\left(\begin{bmatrix} h \\ v \end{bmatrix}\right)\right) \qquad (**)$$

where

$$\begin{aligned}
\alpha\left(\begin{bmatrix} h \\ v \end{bmatrix}, \begin{bmatrix} x \\ u \end{bmatrix}\right) &= h(T)^* P_T x(T) + \int_0^T \begin{bmatrix} h \\ v \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt \\
&= h(T)^* P_T x(T) + \int_0^T h^*(Qx + Su) + v^*(S^* x + Ru)\, dt.
\end{aligned}$$

and (using partial integration)

$$\begin{aligned}
\ell\left(G\left(\begin{bmatrix} h \\ v \end{bmatrix}\right)\right) &= h(0)^* \nu_0 + \int_0^T (Ah + Bv - \dot{h})^* \lambda\, dt \\
&= h(0)^* \nu_0 + \int_0^T (Ah + Bv)^* + h^* \dot{\lambda}\, dt - h(T)^* \lambda(T) + h(0)^* \lambda(0) \\
&= h(0)^*(\nu_0 + \lambda(0)) - h(T)^* \lambda(T) + \int_0^T h^*(A^* \lambda + \dot{\lambda}) + v^*(B^* \lambda)\, dt
\end{aligned}$$

$(**)$ is satisfied if

$$\begin{aligned}
0 &= \nu_0 + \lambda(0) \\
0 &= P_T x(T) - \lambda(T) \\
0 &= A^* \lambda + Qx + Su + \dot{\lambda} \\
0 &= Ru + B^* \lambda + S^* x.
\end{aligned}$$

Basic case: derivation of solution II

The optimality conditions

$$
\begin{aligned}
0 &= \nu_0 + \lambda(0), \\
0 &= P_T x(T) - \lambda(T) \\
0 &= A^*\lambda + Qx + Su + \dot{\lambda} \\
0 &= Ru + B^*\lambda + S^*x.
\end{aligned}
$$

together with the constraints

$$\dot{x} = Ax + Bu, \qquad x(0) = x_0$$

can be written in the nice form

$$\begin{bmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\lambda} \\ \dot{x} \\ \dot{u} \end{bmatrix} = \begin{bmatrix} 0 & A & B \\ A^* & Q & S \\ B^* & S^* & R \end{bmatrix} \begin{bmatrix} \lambda \\ x \\ u \end{bmatrix}, \quad \lambda(T) = P_T\, x(T), \quad x(0) = x_0.$$

The condition $\nu_0 + \lambda(0) = 0$ is trival if λ is determined from the other conditions. From now on, we assume that R **is nonsingular**. Then

$$u = -R^{-1}(B^*\lambda + S^*x).$$

Thus, u can be eliminated from the differential equations, and we obtain

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \underbrace{\begin{bmatrix} \hat{A} & -\hat{R} \\ -\hat{Q} & -\hat{A}^* \end{bmatrix}}_{\mathcal{H}} \begin{bmatrix} x \\ \lambda \end{bmatrix}, \qquad \hat{R} := BR^{-1}B^*, \quad \hat{A} := A - R^{-1}S^*, \quad \hat{Q} := Q - SR^{-1}S^*.$$

$\mathcal{H}$ is a **Hamiltonian matrix:** $\mathcal{H}^*\mathcal{J} = -\mathcal{H}\mathcal{J}$, where $\mathcal{J} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$.

Basic case: derivation of solution III

Boundary value problem again:

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} \hat{A} & -\hat{R} \\ -\hat{Q} & -\hat{A}^* \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix}, \quad \lambda(T) = P_{\scriptscriptstyle T} x(T), \quad x(0) = x_0. \qquad (*)$$

<u>Solution:</u> Consider the fundamental solution

$$\frac{d}{dt} \begin{bmatrix} X_1 & X_2 \\ \Lambda_1 & \Lambda_2 \end{bmatrix} = \begin{bmatrix} \hat{A} & -\hat{R} \\ -\hat{Q} & -\hat{A}^* \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ \Lambda_1 & \Lambda_2 \end{bmatrix}, \qquad \begin{bmatrix} X_1(T) & X_2(T) \\ \Lambda_1(T) & \Lambda_2(T) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

The functions $x = x(t)$, $\lambda = \lambda(t)$ satisfy the first two conditions of $(*)$ iff

$$\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} X_1 & X_2 \\ \Lambda_1 & \Lambda_2 \end{bmatrix} \begin{bmatrix} x(T) \\ \lambda(T) \end{bmatrix} = \begin{bmatrix} X_1 & X_2 \\ \Lambda_1 & \Lambda_2 \end{bmatrix} \begin{bmatrix} x(T) \\ P_{\scriptscriptstyle T}\, x(T) \end{bmatrix} = \begin{bmatrix} X\, x(T) \\ \Lambda\, x(T) \end{bmatrix}$$

where $X := X_1 + P_{\scriptscriptstyle T} X_2$, $\Lambda := \Lambda_1 + P_{\scriptscriptstyle T} \Lambda_2$. Note that

$$\begin{bmatrix} \dot{X} \\ \dot{\Lambda} \end{bmatrix} = \begin{bmatrix} \hat{A} & -\hat{R} \\ -\hat{Q} & -\hat{A}^* \end{bmatrix} \begin{bmatrix} X \\ \Lambda \end{bmatrix}, \qquad \begin{bmatrix} X(T) \\ \Lambda(T) \end{bmatrix} = \begin{bmatrix} I \\ P_{\scriptscriptstyle T} \end{bmatrix}$$

Suppose $X(t)$ is nonsingular for all $t \in [0, T]$. Then the functions λ and x are related via

$$\lambda \;=\; \Lambda\, x(T) = \underbrace{\Lambda X^{-1}}_{=:P} X\, x(T) = P\, x.$$

Differentiating P we obtain

$$\begin{aligned} \dot{P} &= \frac{d}{dt}(\Lambda X^{-1}) = \dot{\Lambda} X^{-1} - \Lambda\, X^{-1} \dot{X} X^{-1} = (-\hat{Q} X - \hat{A}^* \Lambda) X^{-1} - \Lambda X^{-1} (\hat{A} X - \hat{R} \Lambda) X^{-1} \\ &= -(\hat{Q} + \hat{A}^* P + P \hat{A} - P \hat{R} P). \qquad \text{(Riccati-ODE)} \end{aligned}$$

Basic case: derivation of solution IV

The considerations on the foregoing slide suggest the following result, which is easily verified.

Result. Suppose $\hat{J}(h,v) = \alpha(\begin{bmatrix} h \\ v \end{bmatrix}, \begin{bmatrix} h \\ v \end{bmatrix}) \geq 0$ for all solutions (h,v) of $\dot{h} = Ah + bv$, $h(0) = 0$. Suppose further, that R is nonsingular and the Riccati ODE

$$\dot{P} = -(\hat{Q} + \hat{A}^* P + P\hat{A} - P\hat{R}P), \qquad P(T) = P_T$$

has a solution in the interval $[0, T]$. Then the functions x, λ defined by

$$\begin{aligned} \dot{x} &= (\hat{A} - \hat{R}P)x, \qquad x(0) = x_0, \\ \lambda &= Px \end{aligned}$$

solve the boundary value problem

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} \hat{A} & -\hat{R} \\ -\hat{Q} & -\hat{A}^* \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix}, \quad \lambda(T) = P_T x(T), \quad x(0) = x_0.$$

The associated optimal control is given by feedback law

$$u = -R^{-1}(B^* P + S^*)x.$$

The minimum value of $\hat{J}$ is

$$\hat{J}_{min} = -\ell(\begin{bmatrix} x_0 \\ 0 \end{bmatrix}) = -x_0^* \nu_0 = x_0^* \lambda(0) = x_0^* P(0) x_0.$$

What comes next: After treating the general case we give a derivation of the result above without introducing Lagrangian multipliers.

Linear quadratic optimal control: finite time horizon, general case

Linear-quadratic control problem again: Minimize

$$J(x,u) = \underbrace{x(T)^* P_{_T} x(T) + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt}_{\alpha([x,u],[x,u])} - 2\,\Re \underbrace{\left(x(T)^* b_{_T} + \int_0^T \begin{bmatrix} q \\ r \end{bmatrix}^* \begin{bmatrix} x \\ u \end{bmatrix} dt \right)}_{\beta(x,u)}$$

subject to

$$G\left(\begin{bmatrix} x \\ u \end{bmatrix}\right) = \begin{bmatrix} x(0) \\ \Psi_{_T}\, x(T) \\ Ax + Bu - \dot{x} \end{bmatrix} = \begin{bmatrix} x_0 \\ c \\ 0 \end{bmatrix}.$$

Part (3) of (CM)-Proposition: If $G\left(\begin{bmatrix} h \\ v \end{bmatrix}\right) = 0$ implies $\alpha\left(\begin{bmatrix} h \\ v \end{bmatrix}, \begin{bmatrix} h \\ v \end{bmatrix}\right) \geq 0$

and there is a linear functional ℓ such that

$$\alpha\left(\begin{bmatrix} x \\ u \end{bmatrix}, \begin{bmatrix} h \\ v \end{bmatrix}\right) + \ell\left(G\left(\begin{bmatrix} h \\ v \end{bmatrix}\right)\right) = \beta\left(\begin{bmatrix} h \\ v \end{bmatrix}\right)$$

for all piecewise differentiable h and all piecewise continuous v

then (x,u) is a minimizer, and the minimum is $J(x,u) = -\ell(\begin{bmatrix} x_0 \\ c \\ 0 \end{bmatrix})$.

Ansatz: $\ell\left(\begin{bmatrix} w_0 \\ w_T \\ e \end{bmatrix}\right) = \Re\left(w_0^* \nu_0 + w_{_T}^* \nu_{_T} + \int_0^T e^* \lambda \, dt \right)$,

where the vectors ν_0, $\nu_{_T}$ and the continuous and piecewise differentiable function λ (Lagrange multipliers) have to be determined.

The general case: derivation of solution I

We have to satisfy
$$\alpha\left(\begin{bmatrix} h \\ v \end{bmatrix}, \begin{bmatrix} x \\ u \end{bmatrix}\right) + \ell\left(G\left(\begin{bmatrix} h \\ v \end{bmatrix}\right)\right) = \beta\left(\begin{bmatrix} h \\ v \end{bmatrix}\right) \qquad (*)$$

where

$$\begin{aligned}
\alpha\left(\begin{bmatrix} h \\ v \end{bmatrix}, \begin{bmatrix} x \\ u \end{bmatrix}\right) &= h(T)^* P_{T} x(T) + \int_0^T \begin{bmatrix} h \\ v \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt \\
&= h(T)^* P_{T} x(T) + \int_0^T h^*(Qx + Su) + v^*(S^* x + Ru)\, dt, \\
\beta\left(\begin{bmatrix} h \\ v \end{bmatrix}\right) &= \Re\left(h(T)^* b_{T} + \int_0^T h^* q + v^* r\, dt\right)
\end{aligned}$$

and (using partial integration)

$$\begin{aligned}
\ell\left(G\left(\begin{bmatrix} h \\ v \end{bmatrix}\right)\right) &= \Re\left(h(0)^*\nu_0 + (\Psi_{T} h(T))^* \nu_T + \int_0^T (Ah + Bv - \dot{h})^* \lambda\, dt\right) \\
&= \Re\left(h(0)^*\nu_0 + h(T)^* \Psi_{T}^* \nu_T + \int_0^T (Ah + Bv)^* + h^* \dot{\lambda}\, dt - h(T)^*\lambda(T) + h(0)^*\lambda(0)\right) \\
&= \Re\left(h(0)^*(\nu_0 + \lambda(0)) + h(T)^*(\Psi_T^* \nu_T - \lambda(T)) + \int_0^T h^*(A^*\lambda + \dot{\lambda}) + v^*(B^*\lambda)\, dt\right)
\end{aligned}$$

$(*)$ holds if

$$\begin{aligned}
0 &= \nu_0 + \lambda(0) \\
0 &= P_{T} x(T) + \Psi_{T}^* \nu_{T} - \lambda(T) - b_{T} \\
0 &= A^*\lambda + Qx + Su + \dot{\lambda} - q \\
0 &= Ru + B^*\lambda + S^* x - r.
\end{aligned}$$

The general case: derivation of solution II

The optimality conditions

$$\begin{aligned} 0 &= \nu_0 + \lambda(0) \\ 0 &= P_T x(T) + \Psi_T^* \nu_T - \lambda(T) - b_T \\ 0 &= A^*\lambda + Qx + Su + \dot{\lambda} - q \\ 0 &= Ru + B^*\lambda + S^*x - r. \end{aligned}$$

together with the constraints

$$\dot{x} = Ax + Bu, \qquad x(0) = x_0, \qquad \Psi_T x(T) = c$$

can be written in the nice form

$$\begin{bmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\lambda} \\ \dot{x} \\ \dot{u} \end{bmatrix} = \begin{bmatrix} 0 & A & B \\ A^* & Q & S \\ B^* & S^* & R \end{bmatrix} \begin{bmatrix} \lambda \\ x \\ u \end{bmatrix} - \begin{bmatrix} 0 \\ q \\ r \end{bmatrix}, \quad \begin{bmatrix} \lambda(T) \\ c_T \end{bmatrix} = \begin{bmatrix} P_T & \Psi_T^* \\ \Psi_T & 0 \end{bmatrix} \begin{bmatrix} x(T) \\ \nu_T \end{bmatrix}, \quad \begin{bmatrix} \lambda(0) \\ x_0 \end{bmatrix} = \begin{bmatrix} -\nu_0 \\ x(0) \end{bmatrix}.$$

From now on, we assume that R is nonsingular. Then

$$u = -R^{-1}(B^*\lambda + S^*x - r).$$

Thus, u can be eliminated from the differential equations, and we obtain

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \underbrace{\begin{bmatrix} \hat{A} & -\hat{R} \\ -\hat{Q} & -\hat{A}^* \end{bmatrix}}_{\mathcal{H}} \begin{bmatrix} x \\ \lambda \end{bmatrix} + \begin{bmatrix} \hat{r} \\ \hat{q} \end{bmatrix}, \qquad \begin{aligned} &\hat{R} := BR^{-1}B^*, \quad \hat{A} := A - R^{-1}S^*, \quad \hat{Q} := Q - SR^{-1}S^* \\ &\hat{r} = R^{-1}r \qquad\qquad \hat{q} = \text{exercise.} \end{aligned}$$

$\mathcal{H}$ is a Hamiltonian matrix: $\mathcal{H}^*\mathcal{J} = -\mathcal{H}\mathcal{J}$, where $\mathcal{J} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$.

The general case: derivation of solution III

From the foregoing slide we have the following boundary value problem

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} \widehat{A} & -\widehat{R} \\ -\widehat{Q} & -\widehat{A}^* \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} - \begin{bmatrix} \widehat{r} \\ \widehat{q} \end{bmatrix} \qquad \begin{bmatrix} \lambda(T) \\ c_{_T} \end{bmatrix} = \begin{bmatrix} P_{_T} & \Psi^*_{_T} \\ \Psi_{_T} & 0 \end{bmatrix} \begin{bmatrix} x(T) \\ \nu_{_T} \end{bmatrix}, \qquad \begin{bmatrix} \lambda(0) \\ x_0 \end{bmatrix} = \begin{bmatrix} -\nu_0 \\ x(0) \end{bmatrix}.$$

<u>Ansatz for λ:</u> $\boxed{\lambda = Px + g}$ with differentiable functions S, g.

This implies

$$\begin{aligned} \dot{\lambda} &= -\widehat{Q}x - \widehat{A}^*\lambda - \widehat{q} = -\widehat{Q}x - \widehat{A}^*(Px + g) - \widehat{q} = -(\widehat{A}^*P + \widehat{Q})x - (\widehat{A}g + \widehat{q}) \qquad (*) \\ \dot{x} &= \widehat{A}x - \widehat{R}\lambda - \widehat{r} = (\widehat{A} - \widehat{R}P)x - \widehat{R}g - \widehat{r} \qquad (**) \\ \dot{\lambda} &= \frac{d}{dt}(Px + g) = \dot{P}x + P\dot{x} + \dot{g} \\ &= (\dot{P} + P(\widehat{A} - P\widehat{R}P))x + (-P\widehat{R}g - P\widehat{r} + \dot{g}). \qquad (***) \end{aligned}$$

Subtraction of $(*)$ from $(***)$ yields

$$0 = (\dot{P} + \underbrace{P\widehat{A} - P\widehat{R}P + \widehat{Q} + \widehat{A}^*P}_{=:Ric(P)})x + ((\widehat{A}^* - P\widehat{R})g + \widehat{q} - P\widehat{r} + \dot{g})$$

It is easily verified that (x, λ) satisfies the Hamiltonian ODE if

$$\boxed{\dot{P} = -Ric(P), \quad \dot{g} = -(\widehat{A}^* - P\widehat{R})g + P\widehat{r} - \widehat{q}, \quad x \text{ satisfies } (**).}$$

In order to match the end condition we split $g = g_0 + \Psi^*\nu_{_T}$
with an unknown matrix function Ψ.

The general case: derivation of solution IV

Final Ansatz for λ: $\boxed{\lambda = Px + g_0 + \Psi^* \nu_{_T}}$, where

$$\begin{aligned}
\dot{P} &= -Ric(P), \qquad P(T) = P_{_T}, \\
\dot{g}_0 &= -(\widehat{A}^* - P\widehat{R})g_0 + P\widehat{r} - \widehat{q}, \qquad g_0(T) = 0, \\
\dot{\Psi}^* &= -(\widehat{A}^* - P\widehat{R})\Psi^*, \qquad \Psi(T) = \Psi_{_T}.
\end{aligned}$$

Easily verified: If is x defined by

$$\boxed{\dot{x} = (\widehat{A} - \widehat{R}P)x - \widehat{R}(g_0 + \Psi^*\nu_{_T}) - \widehat{r}, \qquad x(0) = x_0,}$$

then (x, λ) solves the Hamiltonian ODE

<u>Needed:</u> $\nu_{_T}$ such that the end condition $\Psi_{_T} x(T) = c$ is matched. We have

$$\frac{d}{dt}\Psi x = \dot{\Psi}x + \Psi\dot{x} = \Psi((\widehat{R}P - \widehat{A})x + \dot{x}) = -\Psi(\widehat{R}(g_0 + \Psi^*\nu_{_T}) + \widehat{r}).$$

$\Rightarrow$

$$\underbrace{\Psi(0)x(0) - \overbrace{\Psi(T)x(T)}^{=c} - \int_0^T \Psi(\widehat{R}g_0 + \widehat{r})\,dt}_{z} = \underbrace{\int_0^T \Psi\widehat{R}\Psi^*\,dt}_{=:W_T}\,\nu_T$$

If w is not in the range of $W_{_T}$ then no $\nu_{_T}$ can be found and the boundary value problem is not solvable. If z is in the range of $W_{_T}$ then a suitable $\nu_{_T}$ can be defined by the Penrose inverse of $W_{_T}$:

$$\nu_T := W_{_T}^+ z.$$

Basic case: addition of differential

Core Lemma (author's terminology): Let

$$J_T^0(x,u) := \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt, \qquad T \in (0,\infty).$$

If $\dot{x} = Ax + Bu$ then for any differentiable function $P(\cdot)$ of Hermitian matrices

$$\begin{aligned} J_T^0(x,u) &= x(0)^*P(0)x(0) - x(T)^*P(T)x(T) \\ &\quad + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \underbrace{\begin{bmatrix} \dot{P} + A^*P + PA + Q & PB + S \\ (PB+S)^* & R \end{bmatrix}}_{=:\mathcal{F}(P)} \begin{bmatrix} x \\ u \end{bmatrix} dt. \end{aligned}$$

Proof. $J_T^0(x,u) = \underbrace{x(0)^*P(0)x(0) - x(T)^*P(T)x(T) + \int_0^T \frac{d}{dt}(x^*Px)\,dt}_{=0} + J_T^0(x,u)$

and

$$\begin{aligned} \frac{d}{dt}(x^*Px) &= x^*\dot{P}x + \dot{x}^*Px + x^*P\dot{x} \\ &= x^*\dot{P}x + (Ax+Bu)^*Px + x^*P(Ax+Bu) = \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} \dot{P} + A^*P + PA & PB \\ (PB)^* & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt \end{aligned}$$

Main Problem in linear quadratic control:
Find P such that $\mathcal{F}(P)$ is positive semidefinite. Why? See following slides.

Remark: P may be constant, i.e. $\dot{P} = 0$. This choice is made if $T = \infty$.

Very important relation to remember

$$\dot{x} = Ax + Bu, \quad P^* = P \qquad \Rightarrow \qquad \tfrac{d}{dt}(x^*Px) = \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} \dot{P} + A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}$$

In the later development P will be constant. Then

$$\frac{d}{dt}(x^*Px) \;=\; \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}.$$

Note that

$$\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} = \begin{bmatrix} A^* & I \\ B^* & 0 \end{bmatrix} \begin{bmatrix} 0 & P \\ P & 0 \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} = \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^* \begin{bmatrix} 0 & P \\ P & 0 \end{bmatrix} \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}$$

This factorization is often used e.g. in papers of Scherer.

Basic case: derivation of solution VI

By the Core Lemma,

$$\underbrace{\int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt}_{J_T^0(x,u)} = x(0)^* P(0)\, x(0) - x(T)^* P(T)\, x(T) + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \underbrace{\begin{bmatrix} \dot{P} + PA + A^*P + Q & PB + S \\ (PB+S)^* & R \end{bmatrix}}_{=:\mathcal{F}(P)} \begin{bmatrix} x \\ u \end{bmatrix} dt.$$

Observations:

(O1) If $\mathcal{F}(P) \geq 0$ then $J_T^0(x,u) \geq x(0)^* P(0)\, x(0) - x(T)^* P(T)\, x(T)$ is bounded below.

(O2) If R is positive definite then (Schur complement)

$$\begin{aligned}
\begin{bmatrix} x \\ u \end{bmatrix}^* \mathcal{F}(S) \begin{bmatrix} x \\ u \end{bmatrix} &= \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} I & 0 \\ R^{-1}(PB+S)^* & I \end{bmatrix} \begin{bmatrix} \dot{P} + Ric(P) & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} I & 0 \\ R^{-1}(PB+S)^* & I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \\
&= x^*(\dot{P} + Ric(P))x + (u + R^{-1}(PB+S)^*x)^* R (u + R^{-1}(PB+S)^*x) \\
&= x^*(\dot{P} + Ric(P))x + \underbrace{v^* R v}_{\geq 0}, \qquad v := u - F\,x, \quad F := -R^{-1}(PB+S)^*,
\end{aligned}$$

where

$$\begin{aligned}
Ric(P) &= A^*P + PA + Q - (PB+S)R^{-1}(PB+S)^* \\
&= P\underbrace{(A - BR^{-1}S^*)}_{=:\hat{A}} + \underbrace{(A - BR^{-1}S^*)^*}_{\hat{A}^*} P + \underbrace{(Q - SR^{-1}S^*)}_{=:\hat{Q}} - P\underbrace{(BR^{-1}B^*)}_{=:\hat{B}} P.
\end{aligned}$$

Simple case: derivation of solution V

From the core lemma and observation (O2) we obtain the following

Theorem. Suppose R is positive definite and the Riccati final value problem

$$\begin{aligned} -\dot{P} &= A^*P + PA + Q - (PB+S)R^{-1}(PB+S)^* \\ &= P\widehat{A} + \widehat{P}^*S + \widehat{Q} - P\widehat{R}P, \qquad P(T) = P_{_T}. \end{aligned}$$

has a solution in $[0,T]$. Then for all (x,u) with

$$\dot{x} = Ax + Bu, \qquad x(0) = x_0$$

the functional

$$\widehat{J}(x,u) = x(T)^* P_{_T} x(T) + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt$$

can be written as

$$\widehat{J}(x,u) = x_0^* P(0) x_0 + \int_0^T \underbrace{v^* R v}_{\geq 0} \, dt, \qquad \text{where} \begin{cases} v := u - F\,x, \\ F := -R^{-1}(PB+S)^*. \end{cases}$$

By definition of v we have, $\dot{x} = (A+BF)x + Bv.$

Obviously, $\widehat{J}$ attains its **minimum** $\widehat{J}_{min} = x_0^* P(0) x_0$ for the control law

$$\begin{aligned} v = 0 \quad \text{i.e.} \quad & u = Fx = -R^{-1}(PB+S)^* x \\ \text{and} \quad & \dot{x}(t) = (A+BF)\,x, \quad x(0) = x_0. \end{aligned}$$

Additional end constraint

Problem: Minimize

$$\hat{J}(x,u) = x(T)^* P_T x(T) + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt \ \text{ subject to } \begin{cases} \dot{x} = A\,x + B\,u, \\ x(0) = x_0, \\ C\,x(T) = c. \end{cases}$$

Theorem. Suppose R is positive definite and the Riccati final value problem

$$-\dot{P} \ = \ A^*P + PA + Q - (PB+S)R^{-1}(PB+S)^*, \qquad P(T) = P_T$$

has a solution in $[0,T]$. Let

$$F = -R^{-1}(PB+S)^*, \qquad W = C\int_0^T \Phi_F(T,s)BR^{-1}B^*\Phi_F(T,t)\,dt\,C^*,$$

where $\Phi_F(t,s)$ is the transition matrix of $A+BF$. Suppose the vector ν satisfies

$$C\,(x_0 - \Phi_F(t,0)\,x_0) = W\,\nu.$$

Then the control

$$u(t) = F\,\Phi_F(t,0)\,x_0 + R^{-1}B^*\Phi_F(T,t)C^*\nu$$

minimzes $\hat{J}$ subject to $C\,x(T) = c$.

Proof. This follows from the representation

$$\hat{J}(x,u) = x_0^* P(0) x_0 + \int_0^T v^*Rv\,dt, \quad u = Fx + v, \quad \dot{x} = Ax + Bv,$$

and our earlier result on minimizing the integral $\int_0^T v^*Rv\,dt$.

Solution of Riccati equation is necessary for minimum

Theorem. Let $R > 0$. Suppose that (A, B) is controllable and there exists $V \in \mathbb{R}$ s.t.

$$V \leq \hat{J}(x,u) = x(T)^* P_{T} x(T) + \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt$$

for all (x, u) such that

$$\dot{x} = A\,x + B\,u,\; x(0) = x_0.$$

Then the Riccati final value problem

$$-\dot{P} \;=\; A^*P + PA + Q - (PB + S)R^{-1}(PB + S)^*, \qquad P(T) = P_{T}$$

has a solution in $[0, T]$.

<u>Proof sketch.</u> The solution exists in $[\theta, T]$ for small $T - \theta$ (by the local existence theorem for ODE of Picard-Lindelöff). Moreover, for any x_θ, $x_\theta^* P(\theta) x_\theta$ minimizes

$$\hat{J}_\theta(x,u) = x(T)^* P_{T} x(T) + \int_\theta^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt \quad \text{subject to } \dot{x} = A\,x + B\,u,\; x(\theta) = x_\theta.$$

Take u which steers x_0 to a given x_θ and then minimizes J_θ for initial value x_θ. Then

$$V \leq \hat{J}(x,u) = \int_0^\theta \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt + \hat{J}_\theta(x,u) \;\leq\; \underbrace{\int_0^\theta \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt}_{a(\theta)} + x_\theta^* P(\theta) x_\theta.$$

Now show that $a(\theta) = \mathcal{O}(\|x_\theta - x_0\|)$. Then it follows that $x_\theta^* P(\theta) x_\theta$ is bounded below. It is also bounded above (because it's a minimum). Thus $P(\theta)$ exists for $\theta \geq 0$.
Full argument in Bell, Jacobson: Singular optimal control.

Example 1 of singular Problem (R not pos. def.)

Special case $R = 0$ (cheap control)

Problem: minimize

$$J(x,u) = \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt = \int_0^T \|x\|_2^2 \, dt \geq 0,$$

subject to

$$\dot{x} = -x + u, \qquad x(0) = x_0.$$

The associated matrix pair $(A, B) = (-I, I)$ is controllable.

Integral can be made arbitrarily small by steering x to 0 at small positive time ϵ. If 0 is reached set $u = 0$ and state will stay at 0 for $t \geq \epsilon$. Conclusion: J can have arbitrarily small values, but 0 (the infimum) is not attained.

$\Rightarrow$ no minimum.

Example 2 of singular Problem (R not pos. def.)

Special case $R = 0$ (cheap control)

Problem: minimize

$$J(x,u) = \int_0^T \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \, dt = \int_0^T 2\,\Re(x^* u)\, dt,$$

subject to

$$\dot{x} = -x + u, \qquad x(0) = x_0.$$

J can be made arbitrarily large and negative by choosing u.

$\Rightarrow$ no minimum.

Homework: Is J bounded below it we add an end constraint $x(T) = x_T$?

Linear Quadratic Optimal Control

Infinite Time Horizon

Kalman-Yakubovich-Popov Lemma

Infinite time horizon I

Problem: Minimize

$$J(x,u)=\int_0^\infty \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt$$

subject to

$$\dot{x}=Ax+Bu, \quad x(0)=x_0, \quad \lim_{t\to\infty}(x(t),u(t))=0, \qquad (x,u)\in\mathcal{L}^2[0,\infty). \qquad (*)$$

where all matrices are constant, and $(x,u)\in\mathcal{L}^2[0,\infty)$ means that

$$\int_0^\infty \|x(t)\|^2\,dt<\infty,\ \int_0^\infty \|u(t)\|^2\,dt<\infty.$$

Further assumption:

(A,B) is completely controllable.

This implies that for each initial value x_0 there are trajectories s.t. $\lim_{t\to\infty} x(t)=0$.

We seek for a minimum of a quadratic functional subject to a linear constraint.

Thus, by our general considerations on linear quadratic minimization:

If for all x_0,

$$-\infty<V_+(x_0):=\inf\{J(x,u)\mid (x.u) \text{ satisfies } (*)\ \}$$

then

there exists a Hermitian P_+ s.t. $V_+(x_0)=x_0^*P_+x_0$ for all x_0.

The basic LMI (linear matrix inequality) and Lur'e equations I

The following is a variant of the Core Lemma.

Suppose there exists a Hermitian P s.t. the following LMI holds.

$$\mathcal{F}(P) := \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & BP \\ PB^* & 0 \end{bmatrix} \geq 0 \quad \text{(positive semidefiniteness)}$$

Then (as for any positive semidefinite matrix) there exists L, K s.t.

$$\mathcal{F}(P) = \begin{bmatrix} K^* \\ L^* \end{bmatrix} \begin{bmatrix} K & L \end{bmatrix}, \qquad \sharp\text{rows}\left(\begin{bmatrix} K & L \end{bmatrix}\right) = \operatorname{rank} \mathcal{F}(P).$$

Componentwise:

$$Q + A^*P + PA = K^*K, \quad S + BP = K^*L, \quad R = L^*L. \qquad \textbf{(Lur'e equations)}.$$

Modified Core Lemma. If $\mathcal{F}(P) \geq 0$ then for $0 \leq t_0 < T$,

$$\begin{aligned} \int_{t_0}^{T} \underbrace{\begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}}_{w(x,u)} dt &= \int_{t_0}^{T} \begin{bmatrix} x \\ u \end{bmatrix}^* \mathcal{F}(P) \begin{bmatrix} x \\ u \end{bmatrix} dt + x(t_0)^* P x(t_0) - x(T)^* P x(T) \\ &= \int_{t_0}^{T} \|Kx + Lu\|^2 \, dt + x(t_0)^* P x(t_0) - x(T)^* P x(T). \end{aligned}$$

The basic LMI (linear matrix inequality) and Lur'e equations II

Modified Core Lemma. If $\mathcal{F}(P) \geq 0$ then for $0 \leq t_0 < T$,

$$\int_{t_0}^{T} \underbrace{\begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}}_{w(x,u)} dt \;=\; \int_{t_0}^{T} \|Kx + Lu\|^2 \, dt + x(t_0)^* P x(t_0) - x(T)^* P x(T).$$

Corollaries.

(1) $\int_{t_0}^{T} w(x.u)\, dt \geq x(t_0)^* P x(t_0) - x(T)^* P x(T)$.

This is a Dissipation inequality. $-w$ is supply for storage $-P$.

(2) Take limits. Set $t_0 = 0$. Then for all u such that $x(T) \to 0$ for $T \to \infty$,

$$J(x,u) = \int_0^{\infty} \|Kx + Lu\|^2 \, dt + x(0)^* P x(0) \;\geq\; x(0)^* P x(0)$$

(3) If $\mathcal{F}(P) \geq 0$ for some P then $J(x,u)$ is bounded below and hence, P_+ exists.

Furthermore, $P_+ \geq P$ since, by def., $x(0)^* P_+ x(0)$ is the greatest lower bound of J.

(4) From (2): If L is nonsingular ($\Leftrightarrow$ R nonsingular) and the control $u = -L^{-1} K x$ is stabilizing ($x(t) \to 0$) then $P = P_+$.

LMI and Riccati inequality

If $R > 0$ we have (Schur complement)

$$\begin{aligned}\mathcal{F}(P) &= \begin{bmatrix} A^*P+PA & BP \\ PB^* & 0 \end{bmatrix} + \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ R^{-1}(PB+S)^* & I \end{bmatrix}^* \begin{bmatrix} Ric(P) & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} I & 0 \\ R^{-1}(PB+S)^* & I \end{bmatrix},\end{aligned}$$

where

$$Ric(P) = A^*P + PA + Q - (PB+S)R^{-1}(PB+S)^*.$$

Thus,

$$\mathcal{F}(P) \geq 0 \quad \Leftrightarrow \quad Ric(P) \geq 0.$$

Suppose $Ric(P) \geq 0$. Then $Ric(P) = K_0^* K_0$ with $\operatorname{rank} K_0 = \operatorname{rank} Ric(P)$.

(K_0 is empty matrix if $Ric(P) = 0$ i.e. if P satisfies Riccati equation). We then have

$$\mathcal{F}(P) = \begin{bmatrix} K_0 & 0 \\ R^{-1/2}(PB+S)^* & R^{-1/2} \end{bmatrix}^* \begin{bmatrix} \underbrace{\begin{matrix} K_0 \\ R^{-1/2}(PB+S)^* \end{matrix}}_{K} & \underbrace{\begin{matrix} 0 \\ R^{-1/2} \end{matrix}}_{L} \end{bmatrix}.$$

Thus,

$$\int_{t_0}^{T} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt = \int_{t_0}^{T} \|Kx + Lu\|^2 \, dt + x(t_0)^* P x(t_0) - x(T)^* P x(T).$$

If $Ric(P) = 0$ and $u = -KL^{-1}x = -R^{-1}(PB+S)^* x$ then the integral vanishes.

If the associated trajectories all satisfy $x(T) \to 0$ as $T \to \infty$ then $P = P_+$.

Equivalence of LMI and dissipation inequality

We have seen on the slide before that

$$\mathcal{F}(P) := \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & BP \\ PB^* & 0 \end{bmatrix} \geq 0, \qquad (*)$$

implies for $t_0 \leq T$ and any trajectory (x, u),

$$\int_{t_0}^{T} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt \geq x(t_0)^* P x(t_0) - x(T)^* P x(T) \qquad (**)$$

Suppose now, that $(**)$ holds. Then, since

$$x(t_0)^* P x(t_0) - x(T)^* P x(T) = -\int_{t_0}^{T} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} PA + A^*P & PB \\ B^*P & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt$$

it follows that for any trajectory

$$f(T) := \int_{t_0}^{T} \begin{bmatrix} x \\ u \end{bmatrix}^* \mathcal{F}(P) \begin{bmatrix} x \\ u \end{bmatrix} dt \geq 0.$$

Differentiating f at $T = t_0$ yields $(*)$.

Minimization problem does not depend on the starting time

Let (x_*, u_*) be a trajectory such that for a given $t_0 \in \mathbb{R}$,

$$\dot{x}_* = Ax_* + Bu_*, \quad x_*(t_0) = x_0, \quad \lim_{t\to\infty}(x_*(t), u_*(t)) = 0, \quad (x_*, u_*) \in \mathcal{L}^2(t_0, \infty) \qquad (*).$$

Define

$$x(t) := x_*(t + t_0), \qquad u(t) := u_*(t + t_0).$$

Then

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \quad \lim_{t\to\infty}(x(t), u(t)) = 0, \quad (x, u) \in \mathcal{L}^2(0, \infty), \qquad (0)$$

and

$$J_*(x_*, u_*) := \int_{t_0}^{\infty} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt = \int_{t}^{\infty} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt = J(x, u).$$

Thus,

$$\begin{aligned} V_*(x_0) \quad &:= \quad \inf \{ J_*(x_*, u_*) \mid (x_*, u_*) \text{ satisfies } (*) \} \\ &= \quad \inf \{ J(x, u) \mid (x, u) \text{ satisfies } (0) \} \\ &= \quad V_+(x_0). \end{aligned}$$

P_+ satisfies dissipation inequality and LMI

For $\epsilon > 0$ and $t_1 \in \mathbb{R}$ there exists a trajectory

$$\dot{x}_* = Ax_* + Bu_*, \quad x_*(t_1) = x_1, \quad \lim_{t\to\infty}(x_*(t), u_*(t)) = 0, \quad (x_*, u_*) \in \mathcal{L}^2(t_1, \infty) \qquad (*)$$

and

$$\int_{t_1}^{\infty} \begin{bmatrix} x_* \\ u_* \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x_* \\ u_* \end{bmatrix} dt \leq V_+(x_1) + \epsilon.$$

Let $t_0 < t_1$. By controllability there exist a trajectory (x, u) which connects (t_0, x_0) with (t_1, x_1). For any such trajectory define (x_c, u_c) by

$$x_c(t) = \begin{cases} x(t) & t \in [t_0, t_1] \\ x_*(t) & t > t_1 \end{cases}, \qquad u_c(t) = \begin{cases} u(t) & t \in [t_0, t_1] \\ u_*(t) & t > t_1 \end{cases}.$$

Then we have for the infima,

$$\begin{aligned} V_+(x_0) &\leq \int_{t_0}^{\infty} \begin{bmatrix} x_c \\ u_c \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x_c \\ u_c \end{bmatrix} dt = \int_{t_0}^{t_1} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt + \int_{t_1}^{\infty} \begin{bmatrix} x_* \\ u_* \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x_* \\ u_* \end{bmatrix} dt \\ &\leq \int_{t_0}^{t_1} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt + V(x_1) + \epsilon. \end{aligned}$$

This holds for any $\epsilon > 0$. Thus,

$$x_0^* P_+ x_0 = V_+(x_0) \leq \int_{t_0}^{t_1} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt + V_+(x_1) = \int_{t_0}^{t_1} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt + x_1^* P_+ x_1.$$

Hence, by our general result before, $\mathcal{F}(P_+) \geq 0$.

Short synopsis

Let (A, B) be controllable. Then the following are equivalent.

(1) The linear matrix inequality

$$\mathcal{F}(P) := \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & BP \\ PB^* & 0 \end{bmatrix} \geq 0 \qquad (*)$$

has a Hermitian solution.

(2) $-\infty < V_+(x_0)$, where

$$V_+(x_0) = \inf \{ J(x,u) \mid \dot{x} = Ax + Bu,\ x(0) = x_0,\ \lim_{t\to\infty} (x(t), u(t)) = 0,\ (x,u) \in \mathcal{L}^2[0,\infty) \}.$$

In this case there exists a maximal solution P_+ of $(*)$, and $V_+(x_0) = x_0^* P_+ x_0$.

If $(*)$ holds for P and $\mathcal{F}(P) = \begin{bmatrix} K^* \\ L^* \end{bmatrix} \begin{bmatrix} K & L \end{bmatrix}$ (Lur'e factorization) then for any $t_0 < T \leq \infty$, and any trajectory connecting (t_0, x_0) with (T, x_T),

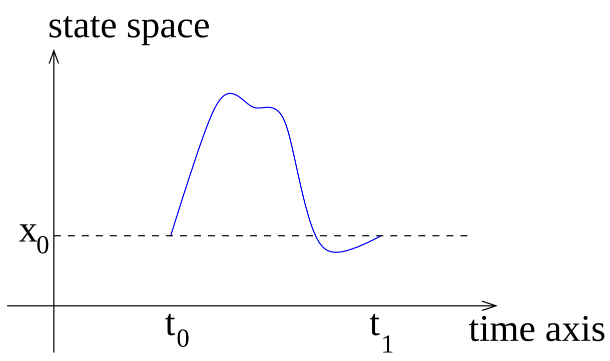
$$\int_{t_0}^{T} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt = \int_{t_0}^{T} \|Kx + Lu\|^2 \, dt + x_0^* P x_0 - x_T^* P x_T \ \geq\ x_0^* P x_0 - x_T^* P x_T.$$

Terminology: The LMI $(*)$ is said to be **feasible** if it has a solution.

Cyclo-dissipativity

The function

$$w(x,u) = \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}$$



is said to be cyclo-dissipative with respect to (A, B) if

for any x_0, any $t_0 < t_1$ and any trajectory $(x.u)$ which connects (t_0, x_0) to (t_1, x_0) (cycle),

$$\int_{t_0}^{t_1} w(x,u)\, dt \geq 0$$

If the LMI $\mathcal{F}(P) \geq 0$ is feasible, then w is cyclo-dissipative since

$$\int_{t_0}^{t_1} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt \;=\; \int_{t_0}^{t_1} \|Kx + Lu\|^2\, dt + \underbrace{x_0^* P x_0 - x_0^* P x_0}_{\text{initial} \,=\, \text{final state} x} \geq 0.$$

On the following page we show that for controllable (A, B),
cyclo-disspativity implies feasibility of the LMI.

Cyclo-dissipativity implies feasibility of LMI (see also picture on next slide)

Suppose

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \quad \lim_{t\to\infty}(x(t), u(t)) = 0, \quad (x, u) \in \mathcal{L}^2(0, \infty).$$

The integral

$$J(x, u) = \int_0^\infty \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt$$

is finite. Thus, for a sufficiently large $T > 0$ the quantities

$$\|x(T)\| \qquad \text{and} \qquad \left| \int_T^\infty \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} \right|$$

are arbitrarily small. By controllability we can modify $u(t)$ for $t > T$ such that the modified control $\tilde{u}$ steers $x(T)$ to $x(T_1) = 0$ and $(\tilde{x}(t), \tilde{u}(t)) = (0, 0)$ for $t \geq T_1 > T$. The difference

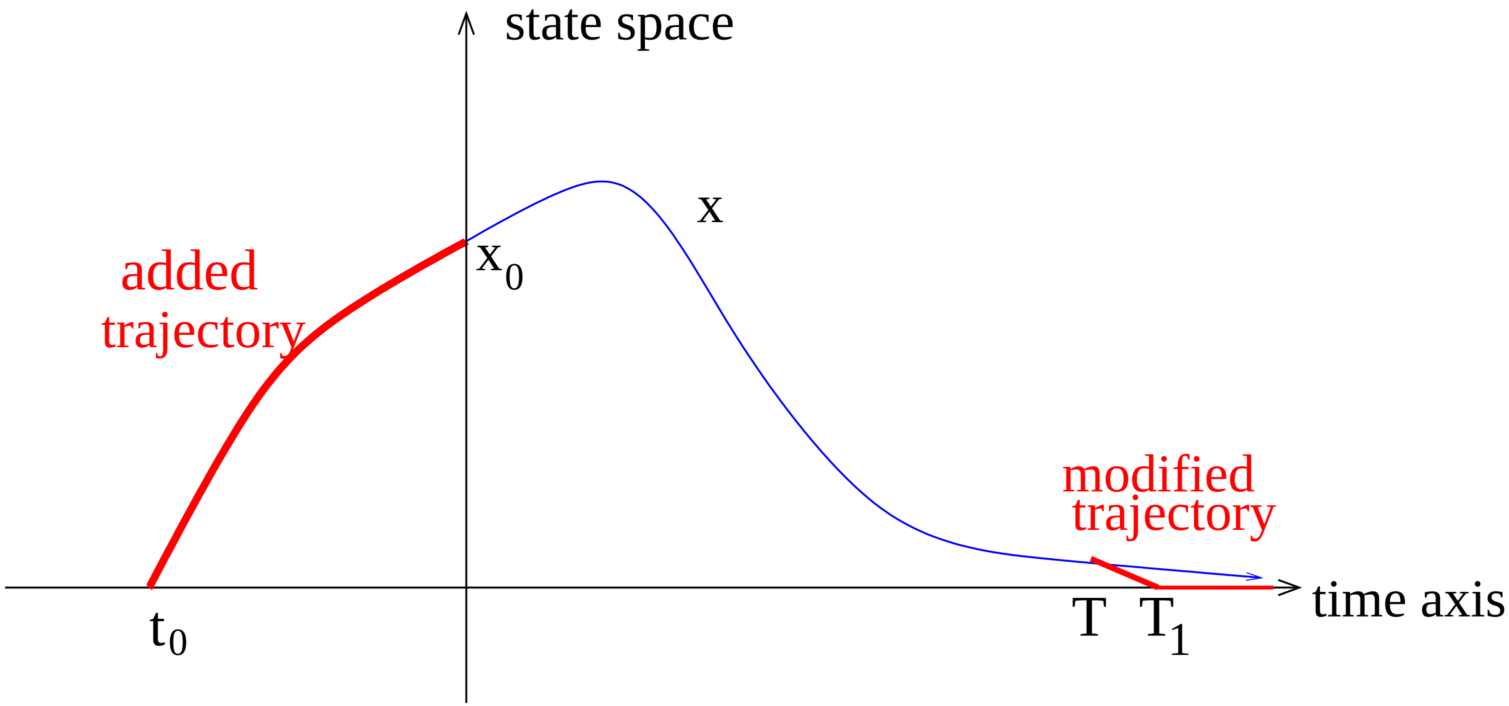
$$\left| \int_T^\infty \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt - \int_T^\infty \begin{bmatrix} \tilde{x} \\ \tilde{u} \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{u} \end{bmatrix} dt \right|, \qquad (\tilde{x}, \tilde{u}) \text{ modified trajectory}$$

can be made arbitrarily small, say $\leq \epsilon$. Now extend $(\tilde{x}, \tilde{u})$ to a trajectory $(\hat{x}, \hat{u})$ connecting $(t_0, 0)$ to $(0, x_0)$ for some $t_0 < 0$. Then $(\hat{x}, \hat{u})$ is a cycle and by cyclo-dissipativity

$$0 \leq \int_{t_0}^\infty \begin{bmatrix} \hat{x} \\ \hat{u} \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{u} \end{bmatrix} dt = \int_{t_0}^0 \begin{bmatrix} \hat{x} \\ \hat{u} \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{u} \end{bmatrix} dt + \int_0^\infty \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt + \epsilon$$

This can be done for any x-trajectory starting at event $(0, x_0) \Rightarrow V_+(x_0) > -\infty \Rightarrow \exists P_+$.

How to form a cycle



Assumption: J(cycle, control of cycle) ≥ 0.

$\Rightarrow \quad J(x,u) \geq -J$(left red trajectory, control of left red trajectory) $- \epsilon$

$\Rightarrow \quad \inf J(x,u) > -\infty \quad \Rightarrow \quad P_+$ exists.

Popov function and Kalman-Yakubovich-Popov Lemma

Popov function:

$$\Pi(s) = \begin{bmatrix} (-\bar{s}\,I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (s\,I - A)^{-1}B \\ I \end{bmatrix}, \qquad s \in \mathbb{C}.$$

Kalman-Yakubovich-Popov Lemma

Suppose (A, B) is completely controllable. Then the following are **equivalent**.

(1) **Popov condition:** For all $\omega \in \mathbb{R}$ such that $\det(i\,\omega\,I - A) \neq 0$,

$$\Pi(i\,\omega) = \begin{bmatrix} (i\,\omega\,I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (i\,\omega\,I - A)^{-1}B \\ I \end{bmatrix} \geq 0.$$

(2) **Feasibility of LMI:** There exists a Hermitian matrix $P = P^*$ such that

$$\mathcal{F}(P) = \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} \geq 0.$$

Supplement (already shown).

(a) If (2) holds then the set of Hermitian P with $\mathcal{F}(P) \geq 0$ has a maximal element P_+ .

(b) If $R > 0$ then

$$\mathcal{F}(P) \geq 0 \quad \Leftrightarrow \quad Q + A^*P + PA - (S + PB)R^{-1}(S + PB)^* \geq 0 \quad \text{(Riccati inequality)},$$

(c) If $R > 0$ and $Q + A^*P + PA - (S + PB)R^{-1}(S + PB)^* = 0$ and $A - BR^{-1}(S + PB)^*$ stable then $P = P_+$.

Feasibility of LMI implies Popov condition (easy direction of equivalence).

We have for any Hermitian P,

$$\begin{bmatrix} (i\,\omega\, I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} \begin{bmatrix} (i\,\omega\, I - A)^{-1}B \\ I \end{bmatrix}$$

$$= \begin{bmatrix} (i\,\omega\, I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} (A - i\omega\, I)^*P + P(A - i\omega\, I) & PB \\ B^*P & 0 \end{bmatrix} \begin{bmatrix} (i\,\omega\, I - A)^{-1}B \\ I \end{bmatrix} = 0.$$

Thus, for any Hermitian P,

$$\Pi(i\omega) = \begin{bmatrix} (i\omega I\, I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (i\omega\, I - A)^{-1}B \\ I \end{bmatrix}$$

$$= \begin{bmatrix} (i\omega I\, I - A)^{-1}B \\ I \end{bmatrix}^* \left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} \right) \begin{bmatrix} (i\omega\, I - A)^{-1}B \\ I \end{bmatrix}$$

(Addition of $\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix}$ does not change the Popov function.) Thus, if

$$\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} \geq 0$$

for some P (feasibility of LMI), then $\Pi(i\omega) \geq 0$ for all $\omega \in \mathbb{R}$ (Popov condition).

Popov condition implies feasibility of LMI (Proof via Fourier transform)

Let $(x.u)$, with $\dot{x} = Ax + Bu$ be a 0-cycle i.e. for some $t_0 < t_1$,

$$(x(t_0), u(t_0)) = (x(t_1), u(t_1)) = 0, \qquad (x(t), u(t)) = (0, 0) \text{ for } t \notin [t_0, t_1].$$

Then the **Fourier transforms** of x, u,

$$\widehat{u}(\omega) = \int_{-\infty}^{\infty} e^{-\omega i t} u(t)\, dt, \qquad \widehat{x}(\omega) = \int_{-\infty}^{\infty} e^{-\omega i t} x(t)\, dt$$

are related by

$$i\omega\, \widehat{x}(\omega) = A\, \widehat{x}(\omega) + B\, \widehat{u}(\omega) \quad \Rightarrow \quad \widehat{x}(\omega) = (i\omega\, I - A)^{-1} B\, \widehat{u}(\omega).$$

By the **Plancherel's theorem (Parseval's Theorem)** and the Popov condition,

$$\int_{t_0}^{t_1} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt = \int_{-\infty}^{\infty} \begin{bmatrix} x \\ u \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \begin{bmatrix} \widehat{x}(\omega) \\ \widehat{u}(\omega) \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} \widehat{x}(\omega) \\ \widehat{u}(\omega) \end{bmatrix} d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \widehat{u}(\omega)^* \underbrace{\begin{bmatrix} (i\,\omega\, I - A)^{-1} B \\ I \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (i\,\omega\, I - A)^{-1} B \\ I \end{bmatrix}}_{\geq 0, \quad \text{Popov condition}} \widehat{u}(\omega)\, d\omega \geq 0.$$

$\Rightarrow$ cyclo-disspativity $\Rightarrow$ LMI-feasibility.

Popov condition implies feasibility of LMI (Proof via convex analysis)

The following proof is due to Anders Rantzer.

Anders Rantzer.
On the Kalman-Yakubovich-Popov lemma.
System & Control Letters 28 (1996), pages 7-10

Rantzer's proof uses the separation theorem for convex sets
and does not require any reasoning about dissipation and cycles.

Reformulation of Popov condition

Popov condition

$$0 \leq \Pi(i\omega) = \begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix} \qquad \text{for all } \omega \in \mathbb{R}$$

can be rewritten as

$$0 \leq w^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} w = \operatorname{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} ww^* \right)$$

for all vectors w such that

$$w = \begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix} u \qquad \text{for some } u \text{ and some } \omega \in \mathbb{R} \qquad (*)$$

Lemma 1 (quantifier elimination). $(*)$ holds if and only if

$$\begin{bmatrix} A^* \\ B^* \end{bmatrix} ww^* \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} ww^* \begin{bmatrix} A & B \end{bmatrix} = 0$$

Proof: Easy. See Rantzer's paper.

Corollary: First reformulation of Popov condition.

Popov condition $\Pi(i\omega) \geq 0$ for all $\omega \in \mathbb{R}$ is equivalent to the statement

$$\operatorname{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} ww^* \right) \geq 0 \qquad \text{for all } w \text{ such that} \qquad \begin{bmatrix} A^* \\ B^* \end{bmatrix} ww^* \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} ww^* \begin{bmatrix} A & B \end{bmatrix} = 0.$$

Rantzer's proof continued

First reformulation of Popov condition (again).

Popov condition $\Pi(i\omega) \geq 0$ for all $\omega \in \mathbb{R}$ is equivalent to the statement

$$\mathrm{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} ww^*\right) \geq 0 \quad \text{for all } w \text{ such that} \quad \begin{bmatrix} A^* \\ B^* \end{bmatrix} ww^* \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} ww^* \begin{bmatrix} A & B \end{bmatrix} = 0.$$

Let W be Hermitian positive semidefinite.
Then W can be written as sum of positive definite rank-1-matrices: $W = \sum_k w_k w_k^*$.

Lemma 2. Suppose $W = \sum_k w_k w_k^*$. Then

$$\begin{bmatrix} A^* \\ B^* \end{bmatrix} W \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} W \begin{bmatrix} A & B \end{bmatrix} = 0 \;\Leftrightarrow\; \begin{bmatrix} A^* \\ B^* \end{bmatrix} w_k w_k^* \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} w_k w_k^* \begin{bmatrix} A & B \end{bmatrix} = 0 \text{ for all } k.$$

Second reformulation of Popov condition

Popov condition $\Pi(i\omega) \geq 0$ for all $\omega \in \mathbb{R}$ is equivalent to the statement

$$\mathrm{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} W\right) \geq 0 \quad \text{for all } W \geq 0 \text{ s.t.} \quad \begin{bmatrix} A^* \\ B^* \end{bmatrix} W \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} W \begin{bmatrix} A & B \end{bmatrix} = 0.$$

Rantzer's proof continued

Second reformulation of Popov condition (again)

Popov condition $\Pi(i\omega) \geq 0$ for all $\omega \in \mathbb{R}$ is equivalent to the statement

$$\operatorname{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} W\right) \geq 0 \quad \text{for all } W \geq 0 \text{ s.t.} \quad \begin{bmatrix} A^* \\ B^* \end{bmatrix} W \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} W \begin{bmatrix} A & B \end{bmatrix} = 0.$$

Third reformulation of Popov condition.

Popov condition $\Pi(i\omega) \geq 0$ for all $\omega \in \mathbb{R}$ holds if and only if the convex sets (cones)

$$K_1 = \left\{ \left(\operatorname{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} W\right), \begin{bmatrix} A^* \\ B^* \end{bmatrix} W \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} W \begin{bmatrix} A & B \end{bmatrix} \right) \;\middle|\; W \geq 0 \right\},$$

$$K_2 = \{ (r, 0) \mid r < 0 \}$$

are disjoint.

Note: K_1, K_2 are convex subsets of $\mathbb{R} \times \mathcal{H}$, where $\mathcal{H}$ is vector space of Herm. matrices.

Separation theorem for convex sets yields:
There exists a nonzero linear functional $f : \mathbb{R} \times \mathcal{H} \to \mathbb{R}$ such that

$$f(K_1) \geq 0, \qquad f(K_2) \leq 0.$$

Functional is of form $\boxed{f(r, H) = \alpha r + \operatorname{tr}(P H), \qquad P \in \mathcal{H},\ \alpha \geq 0.}$

Rantzer's proof continued

For the functional f from the slide before we have for all $W \geq 0$,

$$\begin{aligned}
0 &\leq f\left(\mathrm{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} W\right), \begin{bmatrix} A^* \\ B^* \end{bmatrix} W \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} W \begin{bmatrix} A & B \end{bmatrix}\right) \\
&= \alpha\,\mathrm{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} W\right) + \mathrm{tr}\left(P\left(\begin{bmatrix} A^* \\ B^* \end{bmatrix} W \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} W \begin{bmatrix} A & B \end{bmatrix}\right)\right) \\
&= \alpha\,\mathrm{tr}\left(\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} W\right) + \mathrm{tr}\left(\left(\begin{bmatrix} A^* \\ B^* \end{bmatrix} P \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} P \begin{bmatrix} A & B \end{bmatrix}\right) W\right) \\
&= \mathrm{tr}\left(\left(\alpha \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^* \\ B^* \end{bmatrix} P \begin{bmatrix} I & 0 \end{bmatrix} + \begin{bmatrix} I \\ 0 \end{bmatrix} P \begin{bmatrix} A & B \end{bmatrix}\right) W\right) \\
&= \mathrm{tr}\left(\left(\alpha \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix}\right) W\right)
\end{aligned}$$

This implies

$$0 \leq \alpha \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix}.$$

If $\alpha > 0$ we can replace P by P/α and α by 1, and we are done.

If $\alpha = 0$ then we have

$$0 \leq \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix}.$$

But this contradicts controllablility. See next slide.

Rantzer's proof, last step

Suppose that for Hermitian $P \neq 0$,

$$0 \leq \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix}.$$

This holds iff

$$A^*P + PA \geq 0, \qquad PB = 0. \qquad (*)$$

After a change of coordinates we have $P = \text{diag}(P_1, 0)$ with P_1 nonsingular.

Now write (A, B) in the new coordinates:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

Then

$$A^*P + PA = \begin{bmatrix} A_{11}^* P_1 + P_1 A_{11} & P_1 A_{12} \\ A_{12}^* P_1 & 0 \end{bmatrix}, \qquad PB = \begin{bmatrix} P_1 B_1 \\ 0 \end{bmatrix}.$$

Then $(*)$ yields

$$A = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ B_1 \end{bmatrix}.$$

The subsystem $(A_{11}, 0)$ is not controllable.

Hence, (A, B) is not controllable, contradicting our assumption.

Remark on the controllability assumption

In Rantzer's proof controllability is used to rule out the case $\alpha = 0$.

This is not needed, if K_1 and K_2 are strictly separated,

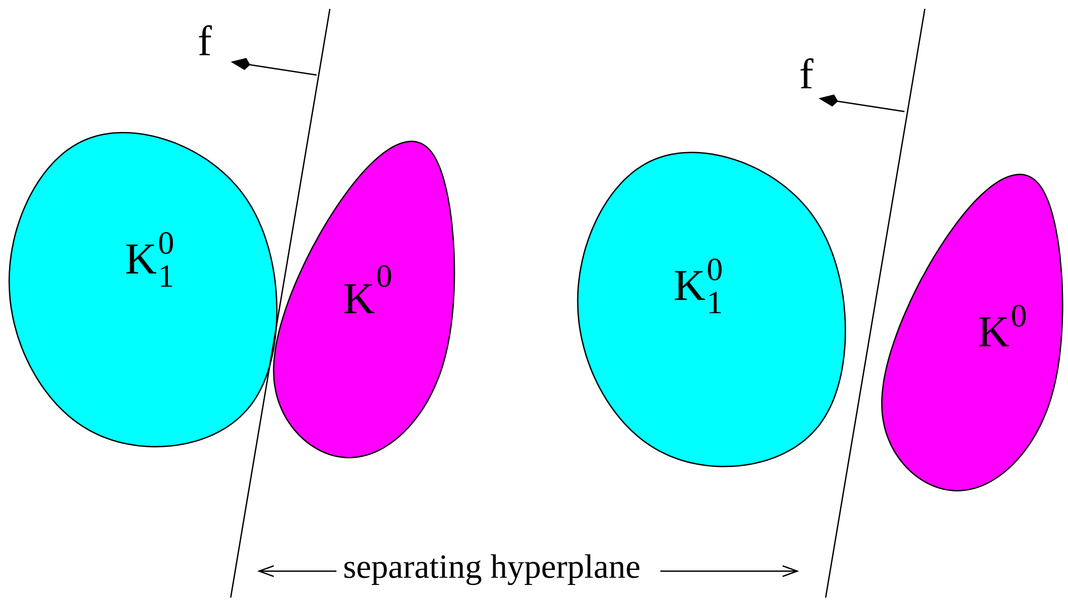
i.e. if there is some space between these sets, since then the separating functional

$$f(r, H) = \alpha\, r + \mathrm{tr}(P\, H)$$

remains separating if α and P are changed a bit.

The following picture illustrates the situation. One has to consider the bounded sets

$$K_j^0 := K_j \cap \{\, W \mid \|W\|_2 = 1\}.$$



This leads to KYP-Lemma for strict inequalities, where now controllability assumption is needed, see next slide.

Kalman-Yakubovich-Popov Lemma for strict inequalities

The following are **equivalent** (no controllability needed).

(1) **Popov condition:** For all $\omega \in \mathbb{R}$ such that $\det(i\omega I - A) \neq 0$,

$$\Pi(i\omega) = \begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (i\omega I - A)^{-1}B \\ I \end{bmatrix} > 0.$$

(2) **Feasibility of LMI:** There exists a Hermitian matrix $P = P^*$ such that

$$\mathcal{F}(P) = \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} + \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} > 0.$$