

Adjoint equations

For the homogeneous system $\dot{x}(t) = A(t)x(t)$, we have here associated adjoint system

$$\dot{p}(t) = -A'(t)p(t).$$

$$\begin{aligned} \text{From } \frac{d}{dt} (p'(t)x(t)) &= \dot{p}'(t)x(t) + p'(t)\dot{x}(t) \\ &= (-A'(t)p(t))'x(t) + p'(t)(A(t)x(t)) \\ &= 0 \end{aligned}$$

it follows that $p'(t)x(t) = p'(t_0)x(t_0) \quad \forall t$.

Moreover, writing $p(t) = \underline{\Phi}_{-A'}(t, t_0)p(t_0)$ we get

$$p'(t_0) \underline{\Phi}'_{-A'}(t, t_0) \underline{\Phi}_A(t, t_0)x(t_0) = p'(t_0)x(t_0) \quad \forall t$$

Since this is true for arbitrary $x(t_0)$, $p(t_0)$ it follows that,

$$\underline{\Phi}'_{-A'}(t, t_0) \underline{\Phi}_A(t, t_0) = \mathbb{1} \quad \forall t$$

$$\Leftrightarrow \underline{\Phi}_{-A'}(t, t_0) = \underline{\Phi}'_A(t_0, t) \quad \forall t.$$

From this we also have a corollary;

$$\begin{aligned} \frac{d}{dt} \underline{\Phi}_A(t_0, t) &= \frac{d}{dt} (\underline{\Phi}'_{-A'}(t, t_0)) \\ &= (-A' \underline{\Phi}_{-A'}(t, t_0))' \\ &= -\underline{\Phi}'_{-A'}(t, t_0) A(t) \\ &= -\underline{\Phi}_A(t_0, t) A(t) \end{aligned}$$

□

Canonical Equations

Consider the linear time-varying system (canonical equations)

$$\begin{pmatrix} \dot{x}(t) \\ \dot{p}(t) \end{pmatrix} = \begin{pmatrix} A(t) & -B(t)B'(t) \\ -L(t) & -A'(t) \end{pmatrix} \begin{pmatrix} x(t) \\ p(t) \end{pmatrix} \quad (C)$$

evolving on \mathbb{R}^{2n} . (Assume $L(t) \equiv L'(t)$). Let $H(t, x, p)$

denote the function

$$H: \mathbb{R} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$$

$$(t, x, p) \mapsto H(t, x, p)$$

$$\begin{aligned} &= \frac{1}{2} x' L(t) x + p' A(t) x \\ &\quad - \frac{1}{2} p' B(t) B'(t) p \end{aligned}$$

Define the gradient of H , $\nabla H = \begin{pmatrix} \frac{\partial H}{\partial x} \\ \frac{\partial H}{\partial p} \end{pmatrix}$.

$$\nabla H = \begin{pmatrix} L(t) & A'(t) \\ A(t) & -B(t)B'(t) \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix}$$

We can then rewrite the given system (C) in the form

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \mathcal{J} \nabla H \quad \text{where } \mathcal{J} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}$$

is a skew-symmetric invertible matrix. Along trajectories of (C) the derivative of H w.r.t. time can be computed using chain rule:

$$\begin{aligned} \frac{dH}{dt} &= \frac{\partial H}{\partial t} + \sum_{i=1}^n \frac{\partial H}{\partial x_i} \dot{x}_i + \sum_{i=1}^n \frac{\partial H}{\partial p_i} \dot{p}_i \\ &= \frac{\partial H}{\partial t} + (\nabla H)' \begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \frac{\partial H}{\partial t} + (\nabla H)' \mathcal{J} \nabla H \end{aligned}$$

Since J is skew, the second term on the right is identically zero. If further, the parameters A, B, L are time-invariant, then $\frac{\partial H}{\partial t} = 0$ and $\frac{dH}{dt} = 0 \Rightarrow H = \text{constant}$.

What are canonical equations (C) good for?

For one thing, solving Riccati equations.

Lemma 1

Let $\Phi(t, t_0)$ denote the $2n \times 2n$ transition matrix for (C). Partition into blocks: $\Phi = \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix}$ of size $n \times n$. Then:

$$(a) \text{ If } \Pi(t, Q, t_1) = \left(\Phi_{22}(t, t_1) - Q \Phi_{12}(t_1, t) \right)^{-1} \left(Q \Phi_{11}(t_1, t) - \Phi_{12}(t_1, t) \right)$$

or equivalently, if

$$(b) \Pi(t, Q, t_1) = \left(\frac{\Phi_{21}(t, t_1) + \Phi_{22}(t_1, t) Q}{\Phi_{11}(t, t_1) + \Phi_{12}(t_1, t) Q} \right)^{-1}$$

then $\Pi(t_1, Q, t_1) = Q$ and, Π satisfies the Riccati equation,

$$\dot{\Pi} = -A' \Pi - \Pi A + \Pi B B' \Pi - L$$

assuming that the indicated inverses exist.

Proof: Let, $[X(t), -P(t)] \triangleq [Q, -1] \begin{bmatrix} \Phi_{11}(t_1, t) & \Phi_{12}(t_1, t) \\ \Phi_{21}(t_1, t) & \Phi_{22}(t_1, t) \end{bmatrix}$

clearly, in part (a) above, the r.h.s = $P^{-1}(t) X(t)$.

To verify that $\Pi(t, Q, t_1) = P^{-1}(t) X(t)$ satisfies the Riccati Equation, differentiate:

$$\frac{d}{dt}(P^{-1}X) = -P^{-1}\dot{P}P^{-1}X + P^{-1}\dot{X}$$

From the definition of $[X, P]$ we see

$$\begin{aligned} [X, -P] &= [Q, -1] \frac{d}{dt} \Phi(t_1, t) \\ &= [Q, -1] \frac{d}{dt} \Phi^{-1}(t, t_1) \\ &= [Q, -1] (-\Phi^{-1}(t, t_1)) \frac{d\Phi^{-1}(t, t_1)}{dt} \Phi^{-1}(t, t_1) \\ &= -[Q, -1] \Phi(t_1, t) \begin{pmatrix} A(t) & -B(t)B'(t) \\ -L(t) & -A'(t) \end{pmatrix} \\ &= -[X, -P] \begin{pmatrix} A & -BB' \\ -L & -A' \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \Rightarrow \dot{X} &= -XA - PL \\ \dot{P} &= -XBB' + PA' \end{aligned}$$

$$\begin{aligned} \Rightarrow \frac{d}{dt}(P^{-1}X) &= -P^{-1}(-XBB' + PA')P^{-1}X + P^{-1}(-XA - PL) \\ &= P^{-1}XBB'P^{-1}X - A'P^{-1}X - P^{-1}XA - L \end{aligned}$$

which is what we set out to prove.

Also $P^{-1}(t)X(t)|_{t=t_1} = Q$ since $\Phi(t_1, t_1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Proof of part (b)

$$\text{Define } \begin{pmatrix} \tilde{X}(t) \\ \tilde{P}(t) \end{pmatrix} = \begin{pmatrix} \Phi_{11}(t, t_1) & \Phi_{12}(t, t_1) \\ \Phi_{21}(t, t_1) & \Phi_{22}(t, t_1) \end{pmatrix} \begin{pmatrix} 1 \\ Q \end{pmatrix}.$$

Clearly, in part (b) above r.h.s = $\tilde{P}(t)\tilde{X}(t)^{-1}$. Rest of the steps similar to the steps in part (a) proof. \square

Remark Matrices of the form

$$P = \begin{pmatrix} A & Q \\ R & -A' \end{pmatrix}$$

where each of the blocks is $n \times n$ and $Q=Q'$, $R=R'$ are called infinitesimally symplectic or hamiltonian matrices.

They satisfy the identity.

$$P'J + JP = \underline{0}$$

Necessary Conditions for Optimality

Theorem For the system $\dot{x}(t) = A(t)x(t) + B(t)u(t)$

$x(t_0) = x_0$, let $u(t)$ be one of the following controls

(a) $u_0(t) = -B'(t) \Phi_A'(t_0, t) \xi$ where ξ satisfies

$$W(t_0, t_1) \xi = x_0 - \Phi_A(t_0, t_1) x_1$$

(b) $u_1(t) = -B'(t) \Pi(t, Q, t_1) x(t)$

$$\dot{\Pi}(t) = -A'(t) \Pi(t, Q, t_1) - \Pi(t, Q, t_1) A(t) - L(t) \\ + \Pi(t, Q, t_1) B(t) B'(t) \Pi(t, Q, t_1)$$

$$\Pi(t_1, Q, t_1) = Q \quad \text{and} \quad L(t) = L'(t) \quad \forall t$$

(c) $u_2(t) = -B'(t) \Pi(t, K_1, t_1) x(t) + v(t)$

$$\dot{\Pi} = -A' \Pi - \Pi A + \Pi B B' \Pi$$

$$\Pi(t_1, K_1, t_1) = K_1$$

and v such that it minimizes

$$\int_{t_0}^{t_1} v(\sigma) v(\sigma) d\sigma \quad \text{for}$$

$$\dot{x}(t) = (A(t) - B(t) B'(t) \Pi(t, K_1, t_1)) x(t) \\ x(t_0) = x_0; \quad x(t_1) = x_1$$

Then, there exists a vector function $p(t)$ (the co-state) such that

$$\begin{pmatrix} \dot{x} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} A & -BB' \\ -L & -A' \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix} \quad x(t_0) = x_0$$

and $u(t) = -B'(t)p(t)$.

~~Proof~~ (a) Let $p(t) = \Phi_A'(t_0, t) \xi$

~~and~~ Then $\dot{p} = -A'(t)p$

with $p_0 = \xi$ (from page 1 of this lecture).

Substituting u_0 in the state equation, we get

$$\dot{x} = Ax - BB'p$$

Picking $L \equiv 0$ ensures that $\dot{p} = -A'p = -Lx - A'p$

This completes the proof of part (a).

(b) Let $p(t) = \Pi(t, Q, t_1)x(t)$

Then substituting u_1 in the state equation we get,

$$\dot{x} = Ax - BB'p \quad \checkmark$$

We need to show

$$\dot{p} = -Lx - A'p$$

Differentiate $\Pi(t, Q, t_1)x(t)$ to get,

$$\dot{p} = \dot{\Pi}x + \Pi\dot{x} = (-A'\Pi - \Pi A - L)x + \Pi BB'\Pi + \Pi(Ax - BB'\Pi x)$$

$$= -A'p - Lx$$

The boundary condition on Π turns into

$$p(t_1) = Qx(t_1).$$

(c) Left as an exercise ▣

Remark We postpone discussion of the infinite horizon optimal control problem and associated algebraic Riccati equations.

Using the Canonical Equations

From proof of part (a) of the Theorem, it is clear that solving (c) for (x_0, p_0) , initial conditions, sweeps out a 'bundle' of state / co-state trajectories as p_0 is varied. Only p_0 s.t. $W(t_0, t_1)p_0 = x_0 - \Phi_A(t_0, t_1)x_1$ will produce trajectory / trajectories satisfying end-point conditions. End-point error associated to a given p_0 can be used to correct p_0 . Similar remarks apply to cases (b) & (c).

Analogues of (c) play a central role in general optimal control problems (not necessarily linear systems with quadratic cost functionals). We will encounter these later.