KTH VETENSKAP OCH KONST

ROYAL INSTITUTE OF TECHNOLOGY

CIAM

The Separation Principle in Stochastic Control, Redux

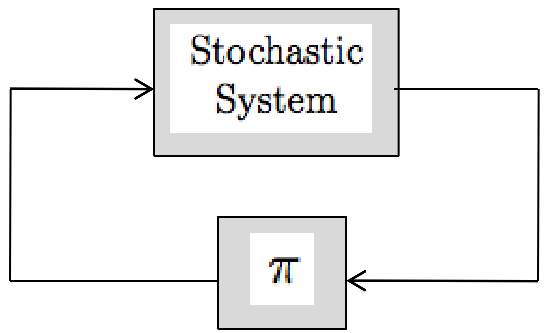
Anders Lindquist
Royal Institute of Technology
Stockholm, Sweden
joining Shanghai Jiao Tong University

- Joint work with Tryphon Georgiou

Swedish-Chinese Control Conference, May, 2011

Stochastic control problem

$$\begin{cases} dx = A(t)x(t)dt + B_1(t)u(t)dt + B_2(t)dw \\ dy = C(t)x(t)dt + D(t)dw \end{cases}$$



Determine control law $\pi : y \mapsto u$ that minimizes

$$J(u) = E\left\{ \int_0^T x(t)'Q(t)x(t)dt + \int_0^T u(t)'R(t)u(t)dt + x(T)'Sx(T) \right\}$$

Separation principle

Under suitable assumptions on the class of admissible control $\pi : y \mapsto u$, the optimal control is

$$u(t) = K(t)\hat{x}(t)$$

where $\hat{x}(t) = E\{x(t) \mid \mathcal{Y}_t\}$,

$$\begin{aligned} d\hat{x} &= A(t)\hat{x}(t)dt + B_1(t)u(t)dt \\ &\quad + L(t)(dy - C(t)\hat{x}(t)dt) \\ \hat{x}(0) &= 0 \end{aligned}$$

with $K(t)$ and $L(t)$ computed via a pair of dual Riccati equations.

$$\boxed{\mathcal{Y}_t := \sigma\{y(s);\, 0 \le s \le t\}}$$

Just assuming that $u(t)$ is $\mathcal{Y}_t$-measurable for each t does not work.

Completion of squares

$$J(u) = E\left\{x(0)'P(0)x(0) + \int_0^T (u - Kx)'R(u - Kx)dt\right\} + \int_0^T \mathrm{tr}(B_2'PB_2)dt$$

where

$$\begin{cases} \dot{P} = -A'P - PA + PB_1R^{-1}B_1'P - Q \\ P(T) = S \end{cases}$$

$$K(t) := -R(t)^{-1}B_1(t)'P(t).$$

With complete state information: $u_{\text{optimal}}(t) = K(t)x(t)$

Incomplete state information

$$E\int_0^T (u-Kx)'R(u-Kx)dt$$

$$= E\int_0^T [(u-K\hat{x})'R(u-K\hat{x})]dt + \text{tr}(K'RK\Sigma)$$

where $\Sigma(t) := E\{\tilde{x}(t)\tilde{x}(t)'\}$.

- Can we conclude that $u = K\hat{x}$ is optimal?

No, not without further assumtions.
Σ may depend on the choice of control.

How?

The tricky point

Due to linearity

$$x(t) = x_0(t) + \int_0^t \Phi(t,s) B_1(s) u(s) ds$$

the control term cancels out:

$$\tilde{x}(t) = \tilde{x}_0(t) := x_0(t) - \hat{x}_0(t),$$

where $\hat{x}_0(t) := E\{x_0(t) \mid \mathcal{Y}_t\}$.

The filtration $\mathcal{Y}_t$, and hence $\hat{x}_0$, might depend on the choice of control u.

We would like to have $\mathcal{Y}_t = \mathcal{Y}_t^0$

Stochastic open loop (SOL): Limit control so as to be adapted to $\{\mathcal{Y}_t^0\}$.

Linear feedback

$$u(t) = u_{\text{deterministic}} + \int_0^t F(t,\tau)dy$$

Then the Gaussian character is preserved, and it can be shown that $\mathcal{Y}_t = \mathcal{Y}_t^0$. Hence,

$$d\tilde{x} = (A - LC)\tilde{x}dt + (B_2 - LD)dw$$
$$\tilde{x}(0) = x(0)$$

$\Sigma(t) := E\{\tilde{x}(t)\tilde{x}(t)'\}$ is independent of u

Lipschitz continuous feedback

Kushner (1967) considered the class of Lipschitz control laws

$$u(t) = \psi(t, \hat{\xi}(t))$$

where

$$\hat{\xi}(t) := E\{x_0(t) \mid \mathcal{Y}_t^0\} + \int_0^t \Phi(t,s)B_1(s)u(s)ds$$

Wonham (1968): If $C(t)$ is square and invertible, then $u(t) = K(t)\hat{x}(t)$ is optimal in the class of Lipschitz control laws of the form

$$u(t) = \psi(t, \hat{x}(t))$$

Fleming & Rishel (1975) removed the assumption on $C(t)$; Lipschitz in y; simpler proof.

Delay in the control action

When $u(t)$ is a function of $y(\tau)$; $0 \leq \tau \leq t - \varepsilon$,

$$\mathcal{Y}_t = \mathcal{Y}_t^0$$

This implies that $\mathcal{Y}_t = \mathcal{Y}_t^0$ always holds in the usual (predictive) discrete-time formulation.

- However, taking $\epsilon \to 0$ and general nonlinear feedback, there is no guarantee that $\mathcal{Y}_t$ is left-continuous
- "Proofs" of the separation theorem (in some popular textbooks) using such limits are circular.

Weak solutions

Davis & Varaya (1972): If we are allowed to consider weak solutions of

$$dx = A(t)x(t)dt + B_1(t)u(t)dt + B_2(t)dw$$

we can change the probability measure (via a Girsanov transformation) so that

$$d\tilde{w} := B_1(t)u(t)dt + B_2(t)dw$$

becomes a new Wiener process, which (under the new probability measure) can be assumed to be unaffected by the control.

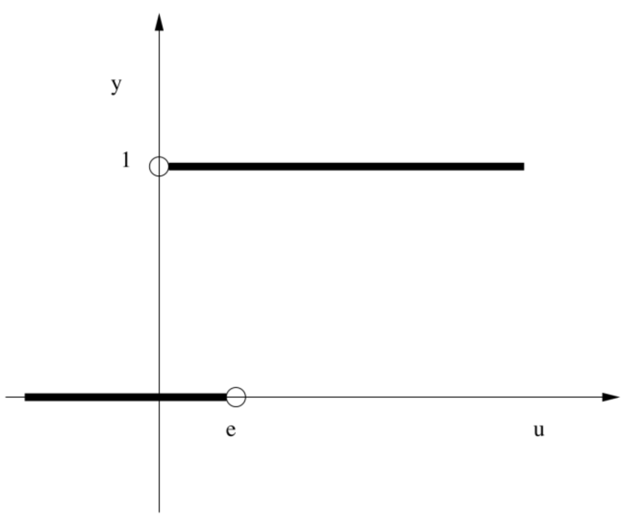
Unclear what is the engineering interpretion of this.

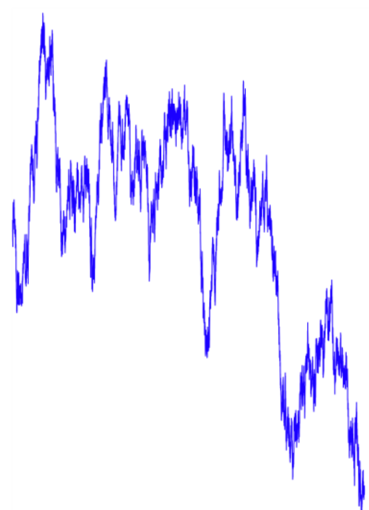
Signals and systems

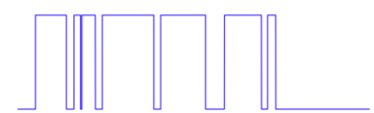
Signals :
sample paths; possibly having bounded discontinuities in D (càdlàg – Skorokhod space)

Systems:
measurable nonanticipatory maps

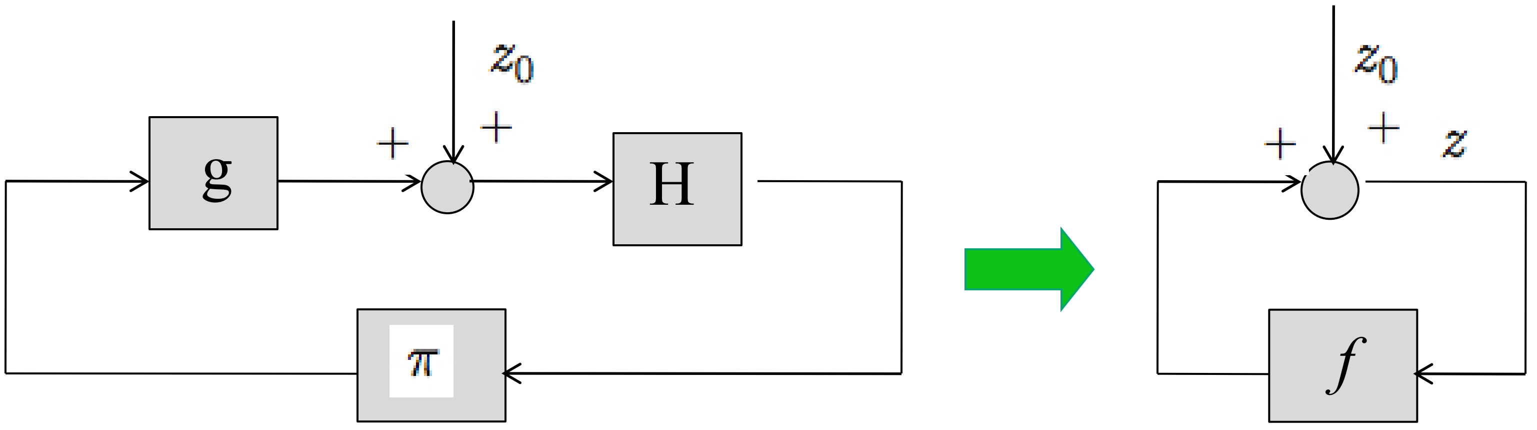
Examples:
i) SDE's that have strong solutions
ii) nonlinearities, hysteresis ($C \to D$), etc.







A general linear stochastic system



$$z(t) = z_0(t) + \int_0^t G(t,\tau)u(\tau)d\tau$$
$$y(t) = Hz(t)$$

$$f = g\pi H$$

where

$$g \; : \; (t,u) \mapsto \int_0^t G(t,\tau)u(\tau)d\tau$$

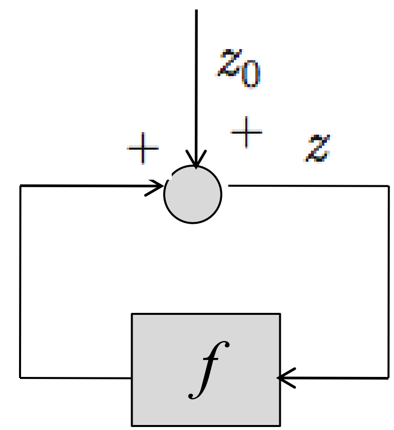
E.g., $z(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$ and $H = [0, I]$

Well-posedness of feedback loop

DEF. A feedback loop

$$z = z_0 + f(z)$$

is *well-posed* if it has a unique solution in D for all $z_0 \in D$ and $(1-f)^{-1}$ is a system.



$(1-f)$ and $(1-f)^{-1}$ systems → $z_0 = z - f(z)$ and $z = (1-f)^{-1} z_0$ → $\mathcal{Z}_t = \mathcal{Z}_t^0$ for $t \in [0, T]$

No new information is created by loop

Complete state information: $\mathcal{Y}_t = \mathcal{Z}_t$; *i.e.*, condition $\mathcal{Y}_t = \mathcal{Y}_t^0$ trivial.

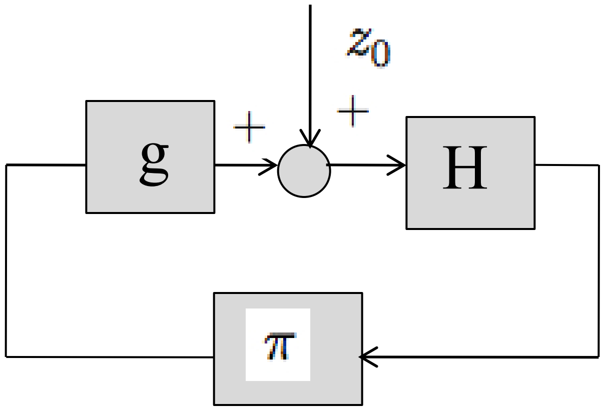
What about incomplete state information?

$$z_1 = \begin{pmatrix} w \\ 0 \end{pmatrix}, \quad z_2 = \begin{pmatrix} 0 \\ w \end{pmatrix}$$ generate the same filtrations, i.e., $\mathcal{Z}_t^{(1)} = \mathcal{Z}_t^{(2)}$

However, for $H = (1 \;\; 0)$,

$$y_1 = (1 \;\; 0)\begin{pmatrix} w \\ 0 \end{pmatrix}, \quad y_2 = (1 \;\; 0)\begin{pmatrix} 0 \\ w \end{pmatrix} \quad \Rightarrow \quad \mathcal{Y}_t^{(1)} \neq \mathcal{Y}_t^{(2)}$$

LEMMA If the feedback loop is well-posed and the readout map H is linear, then $\mathcal{Y}_t = \mathcal{Y}_t^0$ for $t \in [0, T]$.



The separation principle

$$\begin{cases} dx = A(t)x(t)dt + B_1(t)u(t)dt + B_2(t)dw \\ dy = C(t)x(t)dt + D(t)dw \end{cases}$$

Wiener process

$$J(u) = E\left\{\int_0^T x(t)'Q(t)x(t)dt \ + \int_0^T u(t)'R(t)u(t)dt + x(T)'Sx(T)\right\}$$

THEOREM There is a unique $\pi : y \mapsto u$ in the class of well-posed control laws that minimizes $J(u)$, and the optimal control is $u(t) = K(t)\hat{x}(t)$, where $\hat{x}$ is given by the Kalman filter.

The separation principle (general)

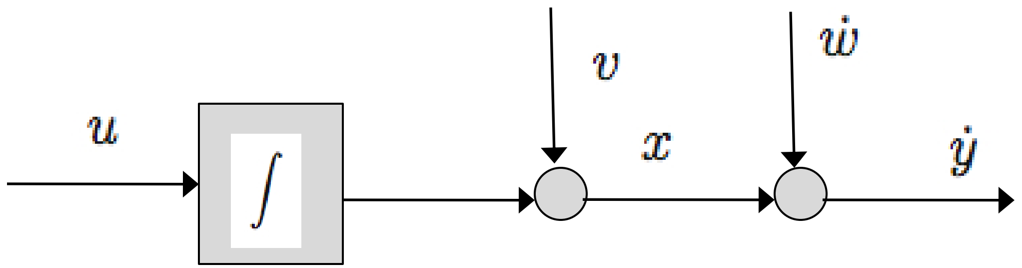
$$\begin{cases} dx = A(t)x(t)dt + B_1(t)u(t)dt + B_2(t)dw \\ dy = C(t)x(t)dt + D(t)dw \end{cases}$$

semimartingale

THEOREM There is a unique $\pi : y \mapsto u$ in the class of well-posed control laws that minimizes $J(u)$, and the optimal control is $u(t) = K(t)\hat{x}(t)$, where $\hat{x}$ is the conditional mean $\hat{x}(t) = E\{x(t) \mid \mathcal{Y}_t\}$.

- strong solutions, but no need for Lipschitz continuity
- signals defined samplewise
- $K(t)$ is still given by a Riccati equation
- in general, the difficult part is constructing $\hat{x}(t) = E\{x(t) \mid \mathcal{Y}_t\}$

Example:
Step change in white noise



$$v(t) = \begin{cases} 1 & t \geq \tau \\ 0 & t < \tau \end{cases} \qquad \text{with } \tau \text{ exponentially distributed}$$

Problem: Find a feedback law $\pi : y \mapsto u$ for the system

$$\begin{cases} dx = u(t)dt + dv, \ x(0) = 0 \\ dy = x(t)dt + dw \end{cases}$$

that minimizes $E\left\{\int_0^T (x^2 + u^2)dt\right\}$.

Example: Solution

- Optimal feedback:

$$u(t) = -p(t)\hat{x}(t)$$

where $\dot{p} = p^2 - 1 \Rightarrow p(t) = \tanh(T - t)$.

- Wonham-Shiryaev filter:

$$d\hat{x} = (1 - \hat{x})dt + udt + \hat{x}(1 - \hat{x})(dy - \hat{x}dt)$$

- Cost: Since $[v, v](t) = v(t)$,

$$\begin{aligned} E\left\{\int_0^T p(t)d[v,v](t)\right\} &= E\left\{\int_\tau^T p(t)dt\right\} \\ &= \ln(\cosh T)(1 - e^{-T}) - \int_0^T \ln(\cosh t)e^{-t}dt. \end{aligned}$$

Key points

- Subject historically marred by incomplete arguments or lengthy derivations
- The natural (engineering) way of thinking about feedback control is in terms of signals
- The natural (and simple) assumption of well-posedness implies that the usual linear feedback control is better than any (reasonable) nonlinear control law
- The Separation Principle holds with semimartingale noise with possible jumps