Estimation, Control, and their Relationship

Robert Bassett

UC Davis Mathematics

September 25th, 2015

Estimation

What is estimation?

- We want a "best guess" of x, an unknown vector.
- Given:
    - Value of z, a related random variable
    - Think of z as a sloppy measurement of z
    - Conditional PDF $f_{X|Z}(x|z)$ of x given z.

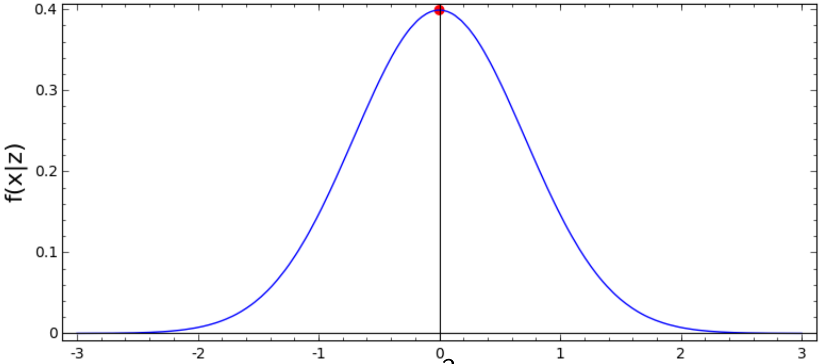
Estimation

By "best guess" we use a maximum likelihood approach.

- We want to find $\hat{x}(z)$ that maximizes

$$\mathbb{P}_{X|Z}(x|z)$$

.



Example

An easy example:

- x is some constant value, and we observe z
- $$z = x + v$$

Example

An easy example:

- x is some constant value, and we observe z
- $$z = x + \underbrace{v}_{\text{Standard normal random variable}}$$

Example

An easy example:

- x is some constant value, and we observe z
- $$z = x + \underbrace{v}_{\text{Standard normal random variable}}$$

By Bayes' theorem,

$$\mathbb{P}_{X|Z}(x|z) = \frac{\mathbb{P}_{Z|X}(z|x) \times \mathbb{P}_X(x)}{\mathbb{P}_Z(z)}$$

So that equivalently, we want to maximize $\mathbb{P}_{Z|X}(z|x)$.

Estimation

Recall that $z = x + v$, so that

$$\mathbb{P}_{Z|X}(z|x) = \mathbb{P}_v(z - x).$$

Estimation

Recall that $z = x + v$, so that

$$\mathbb{P}_{Z|X}(z|x) = \mathbb{P}_v(z - x).$$

Using that v is Gaussian,

$$\mathbb{P}_v(z - x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(z-x)^2}{2}}.$$

This is equivalent to maximizing

$$\frac{-(z - x)^2}{2}$$

Estimation

Recall that $z = x + v$, so that

$$\mathbb{P}_{Z|X}(z|x) = \mathbb{P}_v(z - x).$$

Using that v is Gaussian,

$$\mathbb{P}_v(z - x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(z-x)^2}{2}}.$$

This is equivalent to maximizing

$$\frac{-(z - x)^2}{2}$$

Big Idea

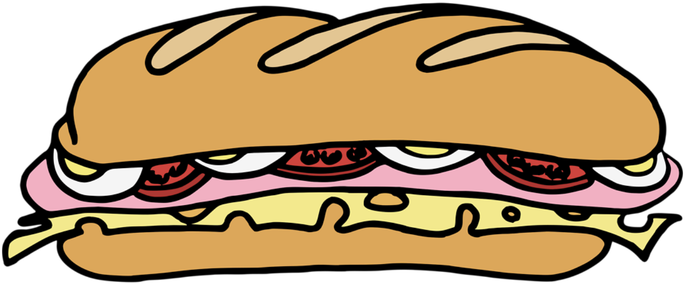
Maximizing $\mathbb{P}_{X|Z}(x|z)$ reduces to minimizing mean-squared error when the measurement corruption is gaussian.

Another example

We want to use sonar to determine the location of a submarine.

Another example

We want to use sonar to determine the location of a submarine.



Another example

We want to use sonar to determine the location of a submarine at time T.

- The submarine moves according to
  $$\underbrace{x(t+1)}_{\text{Position of sub}} = A(t)x(t) + \underbrace{w(t)}_{\text{Gaussian random variable}}.$$

Another example

We want to use sonar to determine the location of a submarine at time T.

- The submarine moves according to
$$\underbrace{x(t+1)}_{\text{Position of sub}} = A(t)x(t) + \underbrace{w(t)}_{\text{Gaussian random variable}}.$$
- Sonar measurements are given by
$$\underbrace{z(t)}_{\text{Sonar Measurement}} = H(t)x(t) + \underbrace{v(t)}_{\text{Gaussian random variable}}.$$

Another example

We want to use sonar to determine the location of a submarine at time T.

- The submarine moves according to
$$\underbrace{x(t+1)}_{\text{Position of sub}} = A(t)x(t) + \underbrace{w(t)}_{\text{Gaussian random variable}}.$$
- Sonar measurements are given by
$$\underbrace{z(t)}_{\text{Sonar Measurement}} = H(t)x(t) + \underbrace{v(t)}_{\text{Gaussian random variable}}.$$
- Assumptions:
  - $\mathbb{E}[v(t)v^T(\tau)] = V(t)\delta_{t,\tau}$
  - $\mathbb{E}[w(t)w^T(\tau)] = R(t)\delta_{t,\tau}$
  - $x(t_0) = x_0$
  - $\mathbb{E}[x(0)x^T(0)] = J_0$.

Another example

Fact of Life

Given measurements $z(0), ..., z(T-1)$, the maximum likelihood estimate of the states $x(0), ..., x(T)$ of the submarine is the minimizer of

$$\frac{1}{2}x^T(0)J_0^{-1}x^T(0) + \sum_{t=0}^{T-1}\frac{1}{2}w^T(t)R^{-1}(t)w(t) \\ + \frac{1}{2}(z(t) - H(t)x(t))^T V^{-1}(t)(z - H(t)x(t))$$

where

$$x(t+1) = A(t)x(t) + w(t), \quad t = 0, .., T-1$$

Another example

Fact of Life

Given measurements $z(0), \ldots, z(T-1)$, the maximum likelihood estimate of the states $x(0), \ldots, x(T)$ of the submarine is the minimizer of

$$\underbrace{\frac{1}{2}x^T(0)J_0^{-1}x^T(0)}_{\text{Initial position}} + \sum_{t=0}^{T-1} \underbrace{\frac{1}{2}w^T(t)R^{-1}(t)w(t)}_{\text{Noise in dynamics}}$$

$$+ \underbrace{\frac{1}{2}(z(t) - H(t)x(t))^T V^{-1}(t)(z - H(t)x(t))}_{\text{Noise in measurements}}$$

where

$$x(t+1) = A(t)x(t) + w(t), \quad t = 0, .., T-1$$

Another example

Fact of Life

Given measurements $z(0), ..., z(T-1)$, the maximum likelihood estimate of the states $x(0), ..., x(T)$ of the submarine is the minimizer of

$$\underbrace{\frac{1}{2}x^T(0)J_0^{-1}x^T(0)}_{\text{Initial position}} + \sum_{t=0}^{T-1} \underbrace{\frac{1}{2}w^T(t)R^{-1}(t)w(t)}_{\text{Noise in dynamics}}$$

$$+ \underbrace{\frac{1}{2}(z(t) - H(t)x(t))^T V^{-1}(t)(z - H(t)x(t))}_{\text{Noise in measurements}}$$

where

$$x(t+1) = A(t)x(t) + w(t), \quad t = 0, .., T-1$$

<u>This is the Kalman Filter problem</u>

Control

What is control?

- Deals with behavior of dynamical systems
- We want to choose inputs to achieve a certain state of the system
- Feedback gives hint to how successful we are
- Adjust inputs, or *controls* accordingly.

Control

What is control?

- Deals with behavior of dynamical systems
- We want to choose inputs to achieve a certain state of the system
- Feedback gives hint to how successful we are
- Adjust inputs, or *controls* accordingly.

Example: A car's cruise control

An Example

Consider a deterministic dynamical system

$$\begin{aligned} x(t+1) &= F(t)x(t) + H(t)u(t) \\ x(0) &= 0 \end{aligned}$$

where u is the control and x is the state.

An Example

Consider a deterministic dynamical system

$$\begin{aligned} x(t+1) &= F(t)x(t) + H(t)u(t) \\ x(0) &= 0 \end{aligned}$$

where u is the control and x is the state.

- Suppose at each time t, we incur a penalty
  - $x^T(t)R(t)x(t) + y^T(t)x(t)$ in the state
  - $u^T(t)V(t)u(t) + z^T(t)u(t)$ in the control
- And then a final penalty $x^T P_T x(T)$

An Example

Control Problem

Find controls $u(0), ..., u(T-1)$ that minimizes

$$\frac{1}{2}x^T(T)P_T x(T) + \sum_{t=0}^{T-1} \frac{1}{2}x^T(t)R(t)x(t) + y^T(t)x(t) + \frac{1}{2}u^T(t)V(t)u(t) + z^T(t)u(t)$$

subject to

$$x(t+1) = F(t)x(t) + H(t)u(t), \quad x(0) = x_0$$

An Example

Control Problem

Find controls $u(0), ..., u(T-1)$ that minimizes

$$\frac{1}{2}x^T(T)P_T x(T) + \sum_{t=0}^{T-1} \frac{1}{2}x^T(t)R(t)x(t) + y^T(t)x(t) + \frac{1}{2}u^T(t)V(t)u(t) + z^T(t)u(t)$$

subject to

$$x(t+1) = F(t)x(t) + H(t)u(t), \quad x(0) = x_0$$

This is the linear-quadratic control problem

The Relationship

What if we construct a dual problem to the linear-quadratic control problem?

The Relationship

What if we construct a dual problem to the linear-quadratic control problem?

We'll use Lagrange duality. Recall

- The primal problem is
$$\min_x \max_y L(x, y)$$
- The dual problem is
$$\max_y \min_x L(x, y)$$

where $L(x, y)$ is a Lagrangian function.

The Relationship

A Lagrangian for the linear-quadratic control problem is given by

$$
\begin{aligned}
L(x, u, \lambda) = \\
&\frac{1}{2}x^T(T)P_T x(T) + \sum_{t=0}^{T-1} \frac{1}{2}x^T(t)R(t)x(t) + y^T(t)x(t) \\
&+ \frac{1}{2}u^T(t)V(t)u(t) + z^T(t)u(t) \\
&+ \underbrace{\lambda(t)^T(x(t+1) - F(t)x(t) - H(t)u(t)) + \lambda(-1)^T x(0)}_{\text{from constraints}}
\end{aligned}
$$

The Relationship

Rewriting

$$
\begin{aligned}
L(x, u, \lambda) = \\
& \frac{1}{2} x^T(T) P_T x(T) + \lambda^T(T-1) x(T) \\
& + \sum_{t=1}^{T-1} \frac{1}{2} x^T(t) R(t) x(t) + \left(y(t) + \lambda(t-1) - F^*(t)\lambda(t)\right)^T x(t) \\
& + \sum_{t=0}^{T-1} \frac{1}{2} u^T(t) V(t) + \left(z(t) - H^*(t)\lambda(t)\right)^T u(t) \\
& + \frac{1}{2} x^T(0) R(0) x(0) + \left(y(0) - F^*(0)\lambda(0) + \lambda(-1)\right)^T x(0)
\end{aligned}
$$

The Relationship

Rewriting

$$L(x,u,\lambda) =$$

$$\frac{1}{2}x^T(T)P_T x(T) + \lambda^T(T-1)x(T)$$

$$+ \sum_{t=1}^{T-1} \frac{1}{2}x^T(t)R(t)x(t) + \Big(\underbrace{y(t) + \lambda(t-1) - F^*(t)\lambda(t)}_{\text{Call this } w(t)}\Big)^T x(t)$$

$$+ \sum_{t=0}^{T-1} \frac{1}{2}u^T(t)V(t) + \Big(\underbrace{z(t) - H^*(t)\lambda(t)}_{\text{Call this } v(t)}\Big)^T u(t)$$

$$+ \frac{1}{2}x^T(0)R(0)x(0) + (y(0) - F^*(0)\lambda(0) + \lambda(-1))^T x(0)$$

The Relationship

A bit of matrix calculus gives us the dual problem

Dual Problem

$$\max_{\lambda} -\frac{1}{2}\lambda^T(T-1)P_T^{-1}\lambda(T-1) - \sum_{t=0}^{T-1}\frac{1}{2}w^T(t)R^{-1}(t)w(t)$$
$$-\sum_{t=0}^{T-1}\frac{1}{2}v^T(t)V^{-1}(t)v(t)$$

subject to $\lambda(t-1) = F^*(t)\lambda(t) + w(t) - y(t), \quad t = 0, .., T-1$

$$z(t) = H^*(t)\lambda(t) + v(t), \quad t = 0, ..., T-1$$

The Relationship

A bit of matrix calculus gives us the dual problem

Dual Problem

$$\max_{\lambda} -\frac{1}{2}\lambda^T(T-1)P_T^{-1}\lambda(T-1) - \sum_{t=0}^{T-1}\frac{1}{2}w^T(t)R^{-1}(t)w(t)$$
$$-\sum_{t=0}^{T-1}\frac{1}{2}v^T(t)V^{-1}(t)v(t)$$

$$\text{subject to } \lambda(t-1) = F^*(t)\lambda(t) + w(t) - y(t), \quad t = 0, .., T-1$$
$$z(t) = H^*(t)\lambda(t) + v(t), \quad t = 0, ..., T-1$$

This is the Kalman Filtering problem!

The Relationship

A bit of matrix calculus gives us the dual problem

Dual Problem

$$\min_{\lambda} \underbrace{\frac{1}{2}\lambda^T(T-1)P_T^{-1}\lambda(T-1)}_{\text{Initial position}} + \sum_{t=0}^{T-1} \underbrace{\frac{1}{2}w^T(t)R^{-1}(t)w(t)}_{\text{Noise in dynamics}}$$

$$+ \sum_{t=0}^{T-1} \underbrace{\frac{1}{2}v^T(t)V^{-1}(t)v(t)}_{\text{Noise in measurements}}$$

$$\text{subject to } \underbrace{\lambda(t-1) = F^*(t)\lambda(t) + w(t) - y(t)}_{\text{Dynamics}}, \quad t = 0, \ldots, T-1$$

$$\underbrace{z(t) = H^*(t)\lambda(t) + v(t)}_{\text{Measurements}}, \quad t = 0, \ldots, T-1$$

The Relationship

Things to note:

- The dynamical system moves *backwards* in time.
- The dynamic noise term now has mean $y(t)$.
- Initial position $\lambda(T-1) \sim \mathcal{N}(0, P_T)$.

The Relationship

Things to note:

- The dynamical system moves *backwards* in time.
- The dynamic noise term now has mean $y(t)$.
- Initial position $\lambda(T-1) \sim \mathcal{N}(0, P_T)$.

The Big Idea

Estimation is Dual to Control

Historical Context

In his seminal 1960 paper, Kalman noticed that Filtering is "dual" to linear-quadratic control.

- By "dual", means have the same solution, running one backwards in time.
- Few authors (none since 1970)* have explored the dual relationship more deeply.

Goals and Further Questions

Goals

- Use existing control and filtering algorithms in a primal-dual relationship to establish better algorithms.
- Apply existing primal-dual algorithms from convex analysis to these problems.
- Use duality to get error bounds for both problems!

Goals and Further Questions

Goals

- Use existing control and filtering algorithms in a primal-dual relationship to establish better algorithms.
- Apply existing primal-dual algorithms from convex analysis to these problems.
- Use duality to get error bounds for both problems!

Further questions

- What assumptions can be removed (Gaussian noise, etc.) while still preserving the duality between estimation and control?
- Can the fact that the dual dynamical system runs backwards in time allow us to use measurements that arrive out of order?

Fin

Thank you for your attention!