Duality in Estimation and Control with Log-Concave Noise

Robert Bassett

Department of Mathematics, University of California Davis
CERMICS Laboratory, École des Ponts ParisTech

June 28th, 2016
ICSP 2016, Buzios
An aircraft flies through the air. Radar stations on the ground take measurements of its location.

- Where is it?
- Where is it going?
- Where has it been?
Estimation Problem

Process:
\[ x_{t+1} = F_t(x_t) + w_t \]

Measurements:
\[ z_t = H_t(x_t) + v_t \]

\(x \in \mathbb{R}^n\), a state vector
\(w \in \mathbb{R}^n\), random noise
\(z \in \mathbb{R}^m\), an observation
\(v \in \mathbb{R}^m\), more noise

\(F\) and \(H\) are vector-valued functions.

Problem
Use measurements \(z\) to construct sequential estimates of states \(x\).
Estimation Problem

Process:

\[ x_{t+1} = F_t(x_t) + w_t \]

Measurements:

\[ z_t = H_t(x_t) + v_t \]

- \( x \in \mathbb{R}^n \), a state vector
- \( w \in \mathbb{R}^n \), random noise
- \( z \in \mathbb{R}^m \), an observation
- \( v \in \mathbb{R}^m \), more noise
- \( H \) and \( F \) are vector-valued functions.
Estimation Problem

Process:
\[ x_{t+1} = F_t(x_t) + w_t \]

Measurements:
\[ z_t = H_t(x_t) + v_t \]

- \( x \in \mathbb{R}^n \), a state vector
- \( w \in \mathbb{R}^n \), random noise
- \( z \in \mathbb{R}^m \), an observation
- \( v \in \mathbb{R}^m \), more noise
- \( H \) and \( F \) are vector-valued functions.

Problem

Use measurements \( z \) to construct sequential estimates of states \( x \).
When dynamics and measurements are **Linear**, and noise terms are **Gaussian**, have the Kalman Filter

- Recursive algorithm for **Maximum-a-Posteriori** (MAP) estimation of system state
  - Mode of the conditional distribution $p(x|z)$
- Uses present measurement, previous state vector and variance matrix
- Fast. Requires only matrix multiplications.
- Ubiquitous in Application
Alternative Methods

What if noise is not Gaussian or the processes are nonlinear?

Sequential Monte Carlo Methods (i.e. Particle Filters)
- Allows for nonlinear dynamics and measurements
- Allows for arbitrary noise distributions
- Monte Carlo technique—slow, convergence results rely on LLN
- Use began in 50’s, rigorous analysis not until ’96!

Extended Kalman Filter
- Estimate noise terms with Normal Distributions
- Linearly estimate dynamics and measurements
Alternative Methods

What if noise is not Gaussian or the processes are nonlinear?

**Sequential Monte Carlo Methods (i.e. Particle Filters)**
- Allows for nonlinear dynamics and measurements
- Allows for arbitrary noise distributions
- Monte Carlo technique—slow, convergence results rely on LLN
- Use began in 50’s, rigorous analysis not until ’96!

**Extended Kalman Filter**
- Estimate noise terms with Normal Distributions
- Linearly estimate dynamics and measurements
A random variable $w$ is log-concave if it has a log-concave density.

$$w \sim e^{-f(x)} \quad \text{where} \quad f : \mathbb{R}^d \rightarrow (-\infty, \infty] \text{ is convex.}$$
A random variable $w$ is log-concave if it has a log-concave density.

$$w \sim e^{-f(x)} \text{ where } f : \mathbb{R}^d \rightarrow (-\infty, \infty] \text{ is convex.}$$

Example of Log-Concave Random Variables

- Normal
- Exponential
- Laplace
- Gamma
- Chi
- Uniform over Convex Set
The Classical Case

Example

Let $w_t$ and $v_t$ be Gaussian.

- $w_t \sim \mathcal{N}(0, Q_t)$
- $v_t \sim \mathcal{N}(0, R_t)$

The MAP Problem is

$$
\min_{x, w} \sum_{t=0}^{T} \frac{1}{2} w'_t Q^{-1}_t w_t + \frac{1}{2} v'_t R^{-1}_t v_t
$$

$$
x_{t+1} = F_t x_t + w_{t+1}, \quad t = 0, ..., T - 1
$$

$$
x_0 = w_0
$$

$$
z_t = H_t x_t + v_t, \quad t = 0, ..., T
$$
The Classical Case

Example

Let \( w_t \) and \( v_t \) be Gaussian.

- \( \mathbf{W}_t \sim \mathcal{N}(0, Q_t) \)
- \( \mathbf{V}_t \sim \mathcal{N}(0, R_t) \)

The MAP Problem is

\[
\min_{x,w} \sum_{t=0}^{T} \frac{1}{2} w_t' Q_t^{-1} w_t + \frac{1}{2} (z_t - H_t x_t)' R_t^{-1} (z_t - H_t x_t) \\
x_{t+1} = F_t x_t + w_{t+1}, \quad t = 0, \ldots, T - 1 \\
x_0 = w_0
\]
The Classical Case

Example

Let \( w_t \) and \( v_t \) be Gaussian.

- \( w_t \sim \mathcal{N}(0, Q_t) \)
- \( v_t \sim \mathcal{N}(0, R_t) \)

The MAP Problem is

\[
\min_{x, w} \sum_{t=0}^{T} \frac{1}{2} w_t' Q_t^{-1} w_t + \frac{1}{2} (z_t - H_t x_t)' R_t^{-1} (z_t - H_t x_t)
\]

- **Dynamic Penalty**

- **Measurement Penalty**

\[
x_{t+1} = F_t x_t + w_{t+1} \quad t = 0, ..., T - 1
\]

\[
x_0 = w_0
\]
Central to Kalman’s derivation is the following theorem:

**Theorem (Kalman, 1960)**

The classical MAP problem is dual to the Linear-Quadratic Regulator problem of optimal control, in the sense that there is a bijection between the Riccati equations that characterize their solutions.
Central to Kalman’s derivation is the following theorem:

**Theorem (Kalman, 1960)**

The classical MAP problem is dual to the Linear-Quadratic Regulator problem of optimal control, in the sense that there is a bijection between the Riccati equations that characterize their solutions.

- A Riccati equation is a matrix equation for $P_t$, where

$$x_t = P_t z_t$$

- Satisfies the conditions for optimality derived from the problem’s Hamiltonian.
- Used to translate problem from one domain (estimation) into another (optimal control).
What is duality?

In these formulations, duality is shown by demonstrating a relationship between the equations that characterize these problems’ solutions.
In these formulations, duality is shown by demonstrating a relationship between the equations that characterize these problems’ solutions. **Is that what we mean by duality?**
Convex Analytic Duality

**Definition (Duality)**

A convex problem \( \inf_{x \in C} f(x) \) is said to be dual to the concave problem \( \sup_{y \in D} g(y) \) if there is a convex-concave function \( L(x; y) \) such that

\[
f(x) = \sup_{y \in D} L(x; y), \quad \text{and} \quad g(y) = \inf_{x \in C} L(x; y)
\]
Definition (Duality)

A convex problem \( \inf_{x \in C} f(x) \) is said to be dual to the concave problem \( \sup_{y \in D} g(y) \) if there is a convex-concave function \( L(x; y) \) such that

\[
  f(x) = \sup_{y \in D} L(x; y), \quad \text{and} \quad g(y) = \inf_{x \in C} L(x; y)
\]

We refer to \( L(x, y) \) as a *saddle function*.

Definition (Strong Duality)

Dual problems are said to satisfy *strong duality* if

\[
  \inf_{x \in C} f(x) = \sup_{y \in D} g(y)
\]
Convex Duality for Classical Case

Theorem (Simon and Stubberud, 1970)

The smoothing problem, with observations $T > t$, given by

\[
\begin{align*}
    x_{t+1} &= F_t x_t + w_t & z_t &= H_t x_t + v_t & x_0 &= w_0 \\
    w_t &\sim \mathcal{N}(0, P_t) & v_t &\sim \mathcal{N}(0, Q_t)
\end{align*}
\]

is dual (in the convex-analytic sense) to the LQR problem

\[
\begin{align*}
    y_{t-1} &= F'_t y_t + H'_t u_t, & y_T &= H'_T u_T
\end{align*}
\]

with cost

\[
\sum_{t=0}^{T} \frac{1}{2} x_t' P_t x_t + \frac{1}{2} u_t' Q_t u_t - z_t' u_t
\]
Log-Concave Case

Let $w_t$ and $v_t$ be log-concave.

- $w_t \sim e^{-f_t(x)}$
- $v_t \sim e^{-g_t(x)}$

The MAP Problem is

$$\min_{x,w} \sum_{t=0}^{T} f(w_t) + g(z_t - H_t x_t)$$

$$x_{t+1} = F_t x_t + w_{t+1} \quad t = 0, \ldots, T - 1$$

$$x_0 = w_0$$
Log-Concave Case

Extension

Let $w_t$ and $v_t$ be log-concave.

- $w_t \sim e^{-f_t(x)}$
- $v_t \sim e^{-g_t(x)}$

The MAP Problem is

$$\min_{x,w} \sum_{t=0}^{T} f_t(w_t) + g_t(z_t - H_t x_t)$$

Dynamic Penalty

Measurement Penalty

$$x_{t+1} = F_t x_t + w_{t+1} \quad t = 0, ..., T - 1$$

$$x_0 = w_0$$
## Estimation and Control Summary

<table>
<thead>
<tr>
<th></th>
<th>Gaussian Noise</th>
<th>Log-Concave</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Convex</strong></td>
<td>Simon and Stubberud (1970)</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>Gaussian Noise</td>
<td>Log-Concave</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td><strong>Convex</strong></td>
<td>Simon and Stubberud (1970)</td>
<td>B. Casey, Wets</td>
</tr>
</tbody>
</table>
**Theorem (B., Casey, Wets 2016)**

A convex analytic dual problem of estimation with log concave noise

\[ w_t \sim e^{-f_t(x)}, \quad v_t \sim e^{-g_t(x)} \]

has as dual the following problem of optimal control

\[
\max_{v_0, \ldots, v_T} \sum_{t=0}^{T} -f_t^*(y_t) - g_t^*(v_t) + z_t v_t \\
\text{subject to } y_t = F'_t y_{t+1} + H'_t v_t, \quad t = 0, \ldots, T - 1 \\
y_T = H'_T v_T
\]

- **Controls** \( v \), **states** \( y \).
- **Dual dynamical system run backwards in time**
Theorem (B., Casey, Wets 2016)

If there exists a pair of sequences \( \{ x_t \}_{t=0}^T \), \( \{ w_t \}_{t=0}^T \) that satisfies \( x_0 = w_0, x_{t+1} = F_t x_t + w_t \) and \( w_t \in \text{int dom } f_t \), \( z_t - H_t x_t \in \text{int dom } g_t \) for each \( t \), then the optimal values of the log-concave estimation problem and the control problem on the previous page are equal.
Theorem (B., Casey, Wets 2016)

If there exists a pair of sequences \( \{x_t\}_{t=0}^T, \{w_t\}_{t=0}^T \) that satisfies \( x_0 = w_0, x_{t+1} = F_t x_t + w_t \) and \( w_t \in \text{int dom } f_t \), \( z_t - H_t x_t \in \text{int dom } g_t \) for each \( t \), then the optimal values of the log-concave estimation problem and the control problem on the previous page are equal.

This is essentially Slater’s Condition

- A constraint qualification
- Guarantees the value function \( \phi(u) = \inf_{x \in X} F(x, u) \) is continuous in a neighborhood of the origin.
An Application

Take: \( g(v_t) = \begin{cases} 
-m_1 v_t & v_t \leq 0 \\
\frac{1}{2} m_2 v_t^2 & v_t > 0 
\end{cases} \)
An Application

The corresponding MAP problem with $w_t \sim N(0, Q)$, $v_t \sim e^{-g(x)}$, $F$ randomly generated and $H$ the sum of components is

$$\min_{x,w} \sum_{t=0}^{T} \frac{1}{2} w_t' Q_t^{-1} w_t + g(z_t - Hx_t)$$

s.t. $x_{t+1} = F_t x_t + w_{t+1}$, $t = 0, ..., T - 1$

$x_0 = w_0$

**Hard to solve** because of the piecewise function $g$. 
Applying this theorem to this example gives the dual as

Minimize over $u, y$:

$$\frac{1}{2}y_t' Q_t y_t + \frac{1}{2} \left( \begin{array}{cc} u_1,t & u_2,t \end{array} \right) \left( \begin{array}{cc} 0 & 0 \\ 0 & m_2 \end{array} \right) \left( \begin{array}{c} u_1,t \\ u_2,1 \end{array} \right) - z_t'(u_1,t + u_2,t)$$

s.t. $y_{t-1} = F'_t y_t + H'_t(u_1,t-1 + u_2,t-1)$

$y_T = H'_T(u_1,T + u_2,T)$

$u_{1,t} \in [-m_1, 0]$  

$u_{2,t} \geq 0$

And we can recover the primal solution $(x^*, w^*)$ from the dual via $w^* = Qy^*$. 
Numerical Application

The diagram illustrates the comparison between the true state and the measured state over time. The true state is represented by blue circles, while the measured state is represented by green circles.

The x-axis represents time, ranging from 0 to 12, and the y-axis represents the state, ranging from -1 to 7. The data points are labeled from 1 to 10, indicating specific timestamps or state values.
Conclusion

1. In certain cases, solve estimation directly from control
2. Simplify optimal control problems with piecewise linear-quadratic penalties (Rockafellar and Wets, 1998)
3. Allows for potential decomposition methods
4. Theoretical link between classical case and log-concave extension
Thank You!

For details see preprint

*Log-Concave Duality in Estimation and Control.* Bassett, Casey, Wets. ’16

at math.ucdavis.edu/~rbassett

Joint work with Michael Casey and Roger Wets
References

- A New Approach to Linear Filtering and Prediction Problems. Kalman. ’60
- Duality of Linear Estimation and Control. Simon and Stubberud.
- General Duality Between Optimal Control and Estimation. Todorov. ’08