Intro to Modeling with Linear Dynamics

MLRG: Nov. 1
Micha Elsner
Jason Pacheco
Problem Statement

- Given a “system” we wish to know the state of the system at any given time.
- Formally, given an initial state vector \( x_0 \in \mathbb{R}^n \) and a time \( t \in T \) we wish to know the state at time \( t \) denoted \( x(t) \).
Why Dynamics?

- Market prices
- Populations
- Moving objects
- Sound waves
- Neural excitation
Example – Double Pendulum
Formal Definition

• A **dynamical system** is a manifold M called the **state space**, and an evolution function \( \Phi: M \rightarrow M \)

• In the double pendulum example we can see that the state space could be as many as 8 dimensions, 6 pos, 2 vel
Simple idea

- We understand discrete models (HMMs) pretty well.
- Let's pretend space is a grid!
- (We have to do that anyway... that's how computers work!)
Simple idea

- How many parameters?
- One probability distribution for each square!
- (Accuracy improves as squares increase...)
- Probably not what we want.

The same trajectory! Why learn it twice?
The Plant Equation

• How do we model dynamical systems?
• The Plant Equation, a.k.a. State Space Model:

  Discrete:
  \[ x(k+1) = F(x(k)) \]

  Continuous:
  \[ \frac{\partial}{\partial t} x(t) = A(x(t)) \]

• As opposed to frequency domain (Laplace)
System Outputs

• The system also produces some output vector,

\[ z(k) = H(x(k)) \]

• Where, 
  \[ H(k) \] is the measurement function.

• We can view \( z(k) \) as a “sample” or a measurement at time \( k \).
Linearity

- The state $x(k)$ is a vector in $\mathbb{R}^d$.
- The transition function tells us about:
  \[ x(t+1) = Ax(t) + b \quad \text{discrete time-invariant} \]
  OR
  \[ \frac{\partial}{\partial t} x(t) = Ax(t) + b \quad \text{continuous time-invariant} \]
- This is linear in that a component of $x(t)$ is a weighted sum of the previous components, $w \cdot x(t-1)$.
- It doesn't mean we move in straight lines.
State Space Model

- As a reiteration, the complete state space model is,
  \[ x(k+1) = Fx(k) + b \]
  \[ z(k) = Hx(k) \]

- This is the discrete time-invariant model, other models include
  - Continuous time-invariant \( \frac{\partial}{\partial t} x(t) = Ax(t) + b \)
  - Discrete time-variant \( x(k+1) = F(k)x(k) + b \)
  - Continuous time-variant \( \frac{\partial}{\partial t} x(t) = A(t)x(t) + b \)
Dynamics

• Still have analogues for everything we could do with HMMs.

Terminology:
Forward algorithm (where am I now, given previous observations?)
  \[ \Rightarrow \text{filtering} \]
Backward algorithm (where did I start, given future observations?)
  \[ \Rightarrow \text{smoothing} \]
Why *Linear* Dynamics?

- There are efficient algorithms (based on the Kalman filter) for linear dynamics and Gaussian noise.
- This isn't always the best choice from a modeling standpoint!
- We'll look at inference later.
Projectile motion

Learned to do this in HS physics.

\[ y(t) = -10t^2 + 100t + 0 \]
Writing the dynamics

- Locally, the function is linear.
- We can write the dynamics as a series of linear differential equations.

\[ y(t) = y(t-1) + y'(t-1) \]
\[ y'(t) = y'(t-1) - 2 \cdot 10 \]

- Matrix form:

\[
\begin{pmatrix}
0 & 100 & -10 \\
1 & 1 & 0 \\
0 & 2 & 1
\end{pmatrix}
\]

state: \( y \, y' \, y'' \)

\( p \, v \, a \)
Observations

- The state \( x(k) \) here is a vector, \((p \; v \; a)\).
- We'll probably only see \( p \).
- In general, the observation can be any (vector-valued) linear function of the state.
- Same as the difference between a Markov model and a Hidden Markov Model.
Observability

- In general, the output $z(k)$ of a system does not necessarily give us the entire state of the system.
- For instance, we don't see instantaneous velocities... only positions.
- A system is **completely observable** if the initial state $x(1)$ can be *fully* and *uniquely* recovered from its output $z(k)$ observed over a finite interval.
Discrete approximation

Iteratively apply the dynamics...

Discretizing time causes errors: 1 step/sec (■).

Smaller steps are better: 100 step/sec (■).

Think “resolution”
Example: Pendulum

• We have the equation of motion for the pendulum,

\[ \ddot{\theta} + \frac{g}{l} \theta = \frac{T_c}{ml^2} \]

• Define state variables:

\[ x_1 = \theta, \ x_2 = \dot{\theta} \]

• Rewrite as two first-order differential eq's:

\[ X_1 = X_2 \]

\[ X_2 = -\frac{g}{l} X_1 + \frac{T_c}{ml^2} \]
Example: Pendulum

- Write differential eq's in state-variable form:

\[
\begin{align*}
\dot{X}_1 &= X_2 \\
\dot{X}_2 &= -\frac{g}{l} X_1 + \frac{T_c}{ml^2}
\end{align*}
\]

- Put in matrix form:

\[
\begin{bmatrix}
\dot{X}_1 \\
\dot{X}_2
\end{bmatrix} =
\begin{bmatrix}
0 & 1 \\
-\frac{g}{l} & 0
\end{bmatrix}
\begin{bmatrix}
X_1 \\
X_2
\end{bmatrix}
+ 
\begin{bmatrix}
0 \\
\frac{1}{ml^2}
\end{bmatrix} T_c
\]
Example: Pendulum

- That takes care of the state, now the output
- We can only observe the angle itself so we have,

\[ z = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \]
Example: Pendulum

\[
\begin{align*}
\dot{\mathbf{x}} &= \begin{bmatrix}
0 & 1 \\
-g/l & 0
\end{bmatrix} \mathbf{x} + \begin{bmatrix}
0 \\
1/ml^2
\end{bmatrix} u \\
\mathbf{z} &= \begin{bmatrix}
1 & 0
\end{bmatrix} \mathbf{x}
\end{align*}
\]
Other functions

Exponential growth:

Equation:
\[ y(t) = 2^{rt} \]

Differential:
\[ y'(t) = r \cdot y(t) \]
(and all the derivatives are the same!)

\[ y(t) = 2r \cdot y(t-1) \]
What we can't do linearly

- Logistic growth:
  \[
  \frac{dP}{dt} = rP \left(1 - \frac{P}{K}\right)
  \]

- Derivative is non-linear.
- Better model of populations:
  - Levels off at carrying capacity.
When we talk about “noise” there are really two types:

- Model noise (gust of wind):
  \[ x(k+1) = Fx(k) + b + v(k) \]

- Measurement Noise (camera shake):
  \[ z(k) = Hx(k) + w(k) \]

Both modeled as additive time-invariant quantities
Gaussian noise

• Easiest noise to work with: additive Gaussian white noise (zero mean).
  \[ x(k+1) \sim N(f(x(k)), \sigma) \]
  \[ x(k+1) = f(x(k)) + v, \ v \sim N(0, \sigma) \]

• Noise is often counted on to absorb non-linearities in the data.
Observation Noise

noise variance = 10
Noise Added to $y$
Noise Added to v
Non-Gaussian Noise

Noise added to positive \( v \), subtracted from negative \( v \)
More Non-Gaussian Noise

Region of high noise variance

Region of low noise variance
Long-term behavior

- Let's consider what happens to an initial state $x$ when we iteratively apply the dynamics.
- It can diverge to $\infty$...
  - As happens with the parabola.
- Or it can converge to a set of states.
  - Like the logistic growth model.
- This set is a *limit set*. 
Stability

- Some limits are stable (*attractors*).
  - Neighboring points converge to the limit.
  - If perturbed, system returns to former equilibrium.
  - \( x(t) = 0.5 \cdot x(t-1) + 1 \), fixed point 2

- A plot like this is a *phase space* diagram...
  - Just states, no \( t \) axis.
Instability

- Not all limit sets are stable.
  - \( x(t) = 2 \cdot x(t-1) - 2 \), fixed point 2

- As time goes by, we expect to find the system near one of its (stable) limits.
- (Equivalent of a stationary distribution for Markov processes)
Traffic

• Lots of models.
  – Most are non-linear (sigmoid acceleration to reach target velocity).

• Some key observations:
  – Simple dynamics lead to complex macro interactions.
  – Small shifts in parameters can cause phase shifts (massive change in macro behavior). One example is introducing trucks into an uphill environment.
  – Three phases: smooth flow, stop+go, jam.
Some video

- Comes from Intelligent Driver Model (IDM)

\[
\dot{v}_\alpha = \frac{dv_\alpha}{dt} = a \left( 1 - \left( \frac{v_\alpha}{v_0} \right)^{\delta} - \left( \frac{s^*(v_\alpha, \Delta v_\alpha)}{s_{\alpha}} \right)^2 \right)
\]

with \( s^*(v_\alpha, \Delta v_\alpha) = s_0 + v_\alpha T + \frac{v_\alpha \Delta v_\alpha}{2 \sqrt{ab}} \)

- Not linear.

- Video from http://www.vwi.tu-dresden.de/~treiber/movie3d/index.html