1 Observation Equations

Symbol	Definition
k	Time Step
i	Channel
s_k^i	Signal during time step k on channel i
μ^i	Mean of s^i
L	constant number of time lags
p_k	intended position
v_k	intended velocity
α_k	intended acceleration
θ_k	parameterization of the system
W_k	covariance of the system parameters
$(\sigma^i)^2$	constant and known variance of each channel (should just be built into Q_k)
H_k	history of signals
Q_k	covariance matrix of the observation noise
D	Number of dimensions (currently only works with 3)
N	Number of of channels
S	Number of parameter sets (ie. position, velocity, acceleration)
A	0 if no affine; 1 if affine used

The signal at channel i during time step k is s_k^i. The N signals have gaussian noise with zero mean and covariance Q. The mean of s_k^i is

$$
\begin{aligned}
\mu^i(k|p_k, v_k, a_k, \theta_k, H_k) = \; & \left.\begin{aligned} & a_0^i(p_0)_x + a_{-1}^i(p_{-1})_x + \ldots + a_{-(L-1)}^i(p_{-(L-1)})_x + \\ & b_0^i(p_0)_y + b_{-1}^i(p_{-1})_y + \ldots + b_{-(L-1)}^i(p_{-(L-1)})_y + \\ & c_0^i(p_0)_z + c_{-1}^i(p_{-1})_z + \ldots + c_{-(L-1)}^i(p_{-(L-1)})_z + \end{aligned}\right\} \quad \text{Position} \\
& \left.\begin{aligned} & d_0^i(v_0)_x + d_{-1}^i(v_{-1})_x + \ldots + d_{-(L-1)}^i(v_{-(L-1)})_x + \\ & e_0^i(v_0)_y + e_{-1}^i(v_{-1})_y + \ldots + e_{-(L-1)}^i(v_{-(L-1)})_y + \\ & f_0^i(v_0)_z + f_{-1}^i(v_{-1})_z + \ldots + f_{-(L-1)}^i(v_{-(L-1)})_z + \end{aligned}\right\} \quad \text{Velocity} \\
& \left.\begin{aligned} & g_0^i(\alpha_0)_x + g_{-1}^i(\alpha_{-1})_x + \ldots + g_{-(L-1)}^i(\alpha_{-(L-1)})_x + \\ & h_0^i(\alpha_0)_y + h_{-1}^i(\alpha_{-1})_y + \ldots + h_{-(L-1)}^i(\alpha_{-(L-1)})_y + \\ & k_0^i(\alpha_0)_z + k_{-1}^i(\alpha_{-1})_z + \ldots + k_{-(L-1)}^i(\alpha_{-(L-1)})_z + \end{aligned}\right\} \quad \text{Acceleration (To implement)} \\
& m\} \quad \text{Affine}
\end{aligned}
$$

2 State Vector

The state vector of the system if the $(DNSL + A + 2D) \times 1$ vector

$$
\left.\begin{array}{r}
\begin{bmatrix} a_0^0 & \cdots & a_{-(L-1)}^0 & b_0^0 & \cdots & b_{-(L-1)}^0 & c_0^0 & \cdots & c_{-(L-1)}^0 \end{bmatrix}^T \oplus \\
\begin{bmatrix} a_0^1 & \cdots & a_{-(L-1)}^1 & b_0^1 & \cdots & b_{-(L-1)}^1 & c_0^1 & \cdots & c_{-(L-1)}^1 \end{bmatrix}^T \oplus \\
\vdots \\
\begin{bmatrix} a_0^{N-1} & \cdots & a_{-(L-1)}^{N-1} & b_0^{N-1} & \cdots & b_{-(L-1)}^{N-1} & c_0^{N-1} & \cdots & c_{-(L-1)}^{N-1} \end{bmatrix}^T \oplus
\end{array}\right\} \text{if using position}
$$

$$
\left.\begin{array}{r}
\begin{bmatrix} d_0^0 & \cdots & d_{-(L-1)}^0 & e_0^0 & \cdots & e_{-(L-1)}^0 & f_0^0 & \cdots & f_{-(L-1)}^0 \end{bmatrix}^T \oplus \\
\begin{bmatrix} d_0^1 & \cdots & d_{-(L-1)}^1 & e_0^1 & \cdots & e_{-(L-1)}^1 & f_0^1 & \cdots & f_{-(L-1)}^1 \end{bmatrix}^T \oplus \\
\vdots \\
\begin{bmatrix} d_0^{N-1} & \cdots & d_{-(L-1)}^{N-1} & e_0^{N-1} & \cdots & e_{-(L-1)}^{N-1} & f_0^{N-1} & \cdots & f_{-(L-1)}^{N-1} \end{bmatrix}^T \oplus
\end{array}\right\} \text{if using velocity}
$$

$$
\left.\begin{array}{r}
\begin{bmatrix} g_0^0 & \cdots & g_{-(L-1)}^0 & h_0^0 & \cdots & h_{-(L-1)}^0 & k_0^0 & \cdots & k_{-(L-1)}^0 \end{bmatrix}^T \oplus \\
\begin{bmatrix} g_0^1 & \cdots & g_{-(L-1)}^1 & h_0^1 & \cdots & h_{-(L-1)}^1 & k_0^1 & \cdots & k_{-(L-1)}^1 \end{bmatrix}^T \oplus \\
\vdots \\
\begin{bmatrix} g_0^{N-1} & \cdots & g_{-(L-1)}^{N-1} & h_0^{N-1} & \cdots & h_{-(L-1)}^{N-1} & k_0^{N-1} & \cdots & k_{-(L-1)}^{N-1} \end{bmatrix}^T \oplus
\end{array}\right\} \text{if using acceleration}
$$

$$
\left.\begin{bmatrix} m \end{bmatrix}^T \oplus \right\} \text{if using affine}
$$

$$
\begin{bmatrix} p_x & p_y & p_z & v_x & v_y & v_z \end{bmatrix}^T
$$

3 State Equation (Prediction)

$$
\begin{aligned}
x_{k|k-1} &= F_{k-1} x_{k-1|k-1} + b_{k-1} \\
u_{k|k-1} &= F_{k-1} W_{k-1|k-1} F_{k-1}^T + Q_{k-1}
\end{aligned}
$$

4 Observation Equation (Update)

$$
\begin{aligned}
(W_{k|k})^{-1} &= (W_{k|k-1})^{-1} + \sum_{i=0}^{N-1} \frac{1}{(\sigma^i)^2} \left[\left(\frac{\partial \mu_k^i}{\partial x_k} \right) \left(\frac{\partial \mu_k^i}{\partial x_k} \right)^T + (-s_k^i + \mu_k^i) \frac{\partial \mu_k^i}{\partial x_k \partial x_k^T} \right]_{x_{k|k-1}} \\
x_{k|k} &= x_{k|k-1} + W_{k|k} \cdot \sum_{i=0}^{N-1} \left[\frac{1}{(\sigma^i)^2} (s_k^i - \mu_k^i) \frac{\partial \mu_k^i}{\partial x_k} \right]_{x_{k|k-1}}
\end{aligned}
$$

5 Matrices for the State and Observation Equations

The filter needs to have F, b, and Q. F is a $(DNSL + A + 2D) \times (DNSL + A + 2D)$ matrix. b is a $(DNSL + A + 2D) \times 1$ matrix. Q is a $(DNSL + A + 2D) \times (DNSL + A + 2D)$ matrix.

The values depend on whether or not the filter is time-variant or time-invariant. Most of the following equation come from section 4.1 of Dynamic Programming and Optimal Control (Volume 1, Third Edition; Bertsekas). Note that some of the notation changes a bit here.

The following linear system is used:

$$x_{k+1} = A_k x_k + B_k u_k + w_k, k = 0, 1, \ldots, N-1$$

x_k is the state, u_k is the control, and w_k is noise. w_k must be zero mean and finite second moment. A_k specifies how the state is updated, and B_k specifies how the control is used.

The quadratic cost is

$$\underset{\substack{w_k \\ k=0,1,\ldots,N-1}}{E} \left\{ x_N' Q_N x_N + \sum_{k=0}^{N-1} (x_k' Q_k x_k + u_k' R_k u_k) \right\}$$

The Q_k must be positive semidefinite symmetric, and R_k must be positive definite symmetric.

The optimal control law for every k is

$$\begin{aligned} \mu_k^*(x_k) &= L_k x_k, \text{ where} \\ L_k &= -(B_k' K_{k+1} B_k + R_k)^{-1} B_k' K_{k+1} A_k \\ K_N &= Q_N \\ K_k &= A_k'(K_{k+1} - K_{k+1} B_k (B_k' K_{k+1} B_k + R_k)^{-1} B_k' K_{k+1}) A_k + Q_k \end{aligned}$$

In the time-invariant case, K satisfies the algebraic Riccati equation

$$K = A'(K - KB(B'KB + R)^{-1} B'K)A + Q$$

The resulting control law is

$$\begin{aligned} \mu^*(x) &= Lx, \text{ where} \\ L &= -(B'KB + R)^{-1} B'KA \end{aligned}$$

K can be found using matrix factorization or by iterating on the Riccati equation until K converges. If A, B, R, and Q are simple, then it might be possible to solve to K explicitly.

6 Things to Try

- Fix position dependent observations.
- Add in affine term and add in acceleration dependent observations.
- Make a general way to add in observations (ex. $\sqrt{v_x}$, $\sin(p_x)$, ...)
- Instead of just having variance of channels, use a covariance.
- Either detect the observation covariance ahead of time or update over time as a state parameter.
- To attenuate the effect that large signals might have on overpowering the signal from other channels, we could switch from a linear firing model to a sigmoid one.