

実用的で簡単なところからスタートして、だんだんと一般的な内容に踏み込んでいく。

1 Luenberger のオブザーバ

現代制御の本だとよく出てくる Luenberger オブザーバが出てくる。こういった構成になっているのか。線形システム

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) \quad (1.1)$$

があったとき、観測値 $y(t)$ は

$$y(t) = Cx(t) \quad (1.2)$$

のように得られる。 $y(t)$ から $x(t)$ を常に得られるとは限らない。このとき最も簡単な方法として思いつくのが、システムと同じものを計算機上にコピーとして用意して制御と同時に計算させる。ただ、モデリングの誤差や初期値の誤差があるために、この方法は理想的すぎる。まずは、よく使われるオブザーバの構成について紹介する。

ちなみにこの観測値と推定値との誤差をとって戻すアイデアは、確率システムで出てくるイノベーション過程の観点からみるとまた一味変わって見える。ざっくりいうと

$$e(t) = y(t) - C\hat{x}(t) \quad (1.3)$$

は観測値から既知の情報分を抜いているので本質的に観測値から得られる新しい情報を意味しているのでイノベーションと呼んでいる。

1.1 一般的なオブザーバの構成

ここで紹介したようなオブザーバだけが許されている訳ではない。オブザーバとして許される、一般的な条件について紹介する。

2 カルマンフィルタ

使い勝手の良さや理解のしやすさから先にオブザーバについて触れたが、この手の手法の理解を進めるためにはカルマンフィルタも触れておいた方がよからうと思う。最終的な構成をみると、ゲインが可変になってるだけでオブザーバとまったく同じやんけ！と思えるので、カルマンフィルタに対する敷居も少し下がる。

2.1 ベイズ推定値から最小分散推定へ

状態 x を推定する手法を $f()$ とするとその誤差 e は

$$e = x - f(y) \quad (2.1)$$

コストを $L(e)$ とすると、コスト関数の期待値は

$$\begin{aligned}
R[f] &= E[L(e)] \\
&= E[L(x - f(y))] \\
&= \int_{-\infty}^{\infty} L(x - f(y))p(x, y)dx dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} L(x - f(y))p(x|y)dx \right) p(y)dy
\end{aligned} \tag{2.2}$$

条件付きベイズリスク

$$R_c[f] = \int_{-\infty}^{\infty} L(x - f(y))p(x|y)dx \tag{2.3}$$

コスト関数の形は他にもあるが、 $L(e) = e^2$ のような 2 乗誤差をとってみると

$$R_c[f] = \int_{-\infty}^{\infty} (x - f(y))(x - f(y))^T p(x|y)dx \tag{2.4}$$

これを最小にすることを考える。簡単のために、スカラーで書く。条件つき期待値を

$$\bar{x}(y) = E[x|y] = \int_{-\infty}^{\infty} xp(x|y)dx \tag{2.5}$$

として

$$\begin{aligned}
R_c[f] &= \int_{-\infty}^{\infty} (x - f(y))^2 p(x|y)dx \\
&= \int_{-\infty}^{\infty} (x + \bar{x}(y) - \bar{x}(y) - f(y))^2 p(x|y)dx \\
&= \int_{-\infty}^{\infty} (x - \bar{x}(y))^2 p(x|y)dx + 2 \int_{-\infty}^{\infty} (x - \bar{x}(y))(\bar{x}(y) - f(y))p(x|y)dx + \int_{-\infty}^{\infty} (\bar{x}(y) - f(y))^2 p(x|y)dx \\
&= \int_{-\infty}^{\infty} (x - \bar{x}(y))^2 p(x|y)dx + 2(\bar{x}(y) - f(y)) \int_{-\infty}^{\infty} (x - \bar{x}(y))p(x|y)dx + (\bar{x}(y) - f(y))^2 \\
&= \int_{-\infty}^{\infty} (x - \bar{x}(y))^2 p(x|y)dx + 2(\bar{x}(y) - f(y)) \left(\int_{-\infty}^{\infty} xp(x|y)dx - \bar{x}(y) \right) + (\bar{x}(y) - f(y))^2 \\
&= E[(x - \bar{x}(y))^2|y] + (\bar{x}(y) - f(y))^2
\end{aligned} \tag{2.6}$$

となるからであり

$$E[(x - \bar{x}(y))^2|y] + (\bar{x}(y) - f(y))^2 \geq E[(x - \bar{x}(y))^2|y] \tag{2.7}$$

推定値を

$$f(y) = \bar{x}(y) \tag{2.8}$$

のように、条件付き期待値としたときに等号成立し最小となる。これを最小分散推定値という。

また、推定誤差 e の期待値は

$$\begin{aligned}
E[e] &= E[x - \bar{x}(y)] \\
&= \int_{-\infty}^{\infty} (x - \bar{x}(y))p(x, y)dx dy \\
&= \int_{-\infty}^{\infty} xp(x, y)dx dy - \int_{-\infty}^{\infty} \bar{x}(y)p(x, y)dx dy \\
&= \int_{-\infty}^{\infty} xp(x, y)dx dy - \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} xp(x|y)dx \right) p(y)dy \\
&= \int_{-\infty}^{\infty} xp(x, y)dx dy - \int_{-\infty}^{\infty} xp(x, y)dx dy \\
&= 0
\end{aligned}$$

なるから、最小分散推定値は不偏推定となる。

2.2 最小分散推定値を時系列に用いる

条件付きベイズリスクを最小にするような推定値は、条件つき期待値で得られることを示した。ここでは確率密度関数の時間的な推移を計算する $p(x_k|Y^k)$ と $p(x_{k+1}|Y^k)$ について紹介する。

観測更新ステップ

$$p(x_k|Y^k) = \frac{p(y_k|x_k)p(x_k|Y^{k-1})}{p(y_k|Y^{k-1})} \quad (2.9)$$

時間更新ステップ

$$p(x_{k+1}|Y^k) = \int_{-\infty}^{\infty} p(x_{t+1}|x_t)p(x_t|Y^t)dx_t \quad (2.10)$$

2.3 線形カルマンフィルタ

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t) + w(t) \quad (2.11)$$

$$y(t) = Cx(t) + v(t) \quad (2.12)$$

雑音はガウシアンを過程する。システム雑音は具体的にはモデルの誤差を想定している。ガウシアンを仮定するのは、ガウシアンが全周波数で同一のパワーを持っているから、ノイズとしては最悪な条件を考えていることになる。やりすぎだと思えば、 $Fv(t)$ とか適当なフィルタをかけてあげればよい。

確率システムからスタートして最適フィルタを構成する。最適というからには評価関数がある。ベイズリスクと呼ばれる推定値との誤差を最小にできる

3 ガウシアンフィルタ

確率密度関数をガウシアンを仮定する。すでに紹介したカルマンフィルタはガウシアンフィルタの線形な場合と言える。非線形な場合はガウシアンになるとは限らないためカルマンフィルタのように条件付き期待値を

求めたとしても不偏推定値とはならない。そういう場合はガウシアンで確率密度関数を表現する混合ガウシアンフィルタもあるが、計算コストとのトレードオフも考えると十分なケースも多い。

ベイズリスクを最小にする推定値は以下の条件付き期待値で

$$x \tag{3.1}$$

3.1 拡張カルマンフィルタ

テイラー展開して線形化して線形カルマンフィルタのアルゴリズムに乗せる。ヤコビアン計算が必要になる。

3.2 Gauss-Hermite Filter

確率密度関数にガウシアンを仮定しているとき、モーメントの計算における積分は Gauss 求積のガウスエルミート型になっている。

3.3 UnscentedKalmanFilter

3.3.1 Unscented 変換

UT の気持ち

$$y = h(x) \tag{3.2}$$

ガウスエルミート求積でもわかるように、適切な重みと点を決めれば積分を近似的に求めることができる。似たようなアイデアで、UT では確率変数の非線形変換における、アンサンブル平均で期待値、分散を求める。具体的には $2n$ 個の sigma point と呼ばれる点を以下のように生成する。

$$x^{(i)} = \bar{x} + x_{sigma}^{(i)} \tag{3.3}$$

$$x_{sigma}^{(i)} = (\sqrt{nP})_i^T \tag{3.4}$$

$$x_{sigma}^{(i+n)} = -(\sqrt{nP})_i^T \tag{3.5}$$

sigma point の変換先は

$$y^{(i)} = h(x^{(i)}) \tag{3.6}$$

その期待値

$$\bar{x}_u = \tag{3.7}$$

$$P_u = \tag{3.8}$$

4 Guassian Sum Filter

非線形な確率密度関数を陽に扱う方法として、ガウシアンとの和で確率密度関数を近似する方法を考える。

5 Maximum Correntropy Kalman Filter

Correntropy は二つの確率変数の類似性測度である。二つの確率変数 X, Y が与えられたとき、結合確率密度を $F(x, y)$ として correntropy $V(x, y)$ は以下で定義される

$$V(x, y) = E[\kappa(X, Y)] = \int \kappa(x, y) dF_{XY}(x, y) \quad (5.1)$$

5.1 Maximum Correntropy 基準

6 拘束条件

6.1 constrained

まず拘束条件なしで得られる推定値

$$\hat{x}_k = \arg \max_{x_k} p(x_k | Y^k) \quad (6.1)$$

を求める。これは条件つき期待値

$$x_k = E(x_k | Y^k) \quad (6.2)$$

例として、拘束条件付きの推定値 \tilde{x}_k としたときこの推定値に拘束条件 $D\tilde{x}_k = d$ をつける。ラグランジアンを以下で定義する。

$$L = (\tilde{x}_k - \hat{x}_k)^T P_k^{-1} (\tilde{x}_k - \hat{x}_k) + 2\lambda (D\tilde{x}_k - d) \quad (6.3)$$

$$\frac{\partial L}{\partial x_k} = P_k^{-1} (\tilde{x}_k - \hat{x}_k) + D^T \lambda = 0 \quad (6.4)$$

$$\frac{\partial L}{\partial \lambda} = D\tilde{x}_k - d = 0 \quad (6.5)$$

上式より推定値 \hat{x}_k から \tilde{x}_k を求める更新則を求める

$$\begin{aligned} -P_k^{-1} (\tilde{x}_k - \hat{x}_k) &= D^T \lambda \\ -(D^T)^{-1} P_k^{-1} (D^{-1} d - \hat{x}_k) &= \lambda \\ -(D^T)^{-1} P_k^{-1} D^{-1} (d - D\hat{x}_k) &= \lambda \\ \lambda &= (DP_k D^T)^{-1} (D\hat{x}_k - d) \end{aligned} \quad (6.6)$$

であるから

$$\begin{aligned}\tilde{x}_k &= \hat{x}_k - P_k D^T \lambda \\ &= \hat{x}_k - P_k D^T (D P_k D^T)^{-1} (D \hat{x}_k - d)\end{aligned}\tag{6.7}$$

これは拘束軌道に接する最適値を求めている。

6.2 truncated

状態変数の値に上限、下限があるときに確率密度関数をそこで打ち切る方法

7 moving horizon estimation