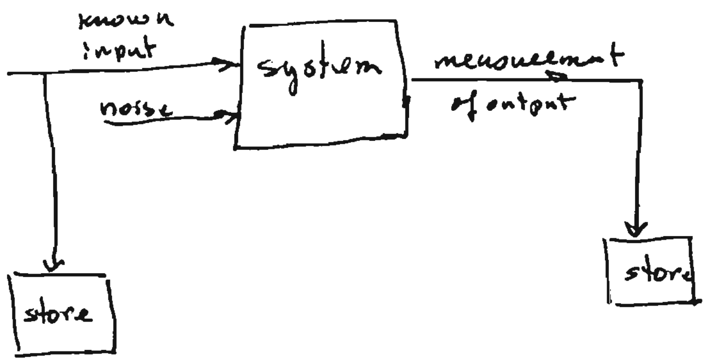
Prediction of Future Behavior

- another approach to the problem (Ljung, Ch. 3)

Given a system



Typically, measured inputs and outputs are stored : $\{(u(t), y(t))\}_{t=0}^{N-1}$

Given these data points, and the next input $u(N)$, what is the best predicted output $y(N)$?

System Model:

$$y(t) = G(q)u(t) + H(q)e(t) \quad \begin{cases} u(t) \text{ known} \\ e(t) \text{ i.i.d. noise} \end{cases}$$

where

$$G(q) = \sum_{k=1}^{\infty} g(k) q^{-k}$$

$$H(q) = 1 + \sum_{k=1}^{\infty} h(k) q^{-k}$$

What about prediction m time steps into the future ?

Suppose first that $u \equiv 0$. Let

$$v(t) = H(q)\,e(t) = \sum_{k=0}^{\infty} h(k)\,e(t-k) \qquad (h(0) = 1.)$$

where H is stable $\left(\sum_{k=0}^{\infty} |h(k)| < \infty \right)$.

Given $\{v(t)\}_{t=0}^{T}$, when can the input noise process be recovered?

Remember that $H(q)$ can be viewed as $H(z)$ - the transfer function from e to v. In this case, $H^{-1}(z)$ exists and is stable $\iff$ $H(z)$ has no zeros in $\{ z \mid |z| \geq 1 \}$. When $H^{-1}(z)$ exists, it can be written as

$$\frac{1}{H(z)} = \sum_{k=0}^{\infty} \tilde{h}(k)\, z^{-k}$$

and it is stable $\iff$ $\sum_{k=0}^{\infty} |\tilde{h}(k)| < \infty$.

In this case, $e(t) = H^{-1}(q)\,v(t)$ where

$$H^{-1}(q) = \sum_{k=0}^{\infty} \tilde{h}(k)\, q^{-k}$$

Example: Moving Average of Noise

$$v(t) = e(t) + c e(t-1)$$

$$H(z) = 1 + cz^{-1} = \frac{z+c}{z}$$

$$H^{-1}(z) = \frac{1}{1+cz^{-1}} = \sum_{k=0}^{\infty} (-c)^k z^{-k} \qquad \text{where } |c| < 1.$$

Therefore, to recover $e(t)$:

$$e(t) = \sum_{k=0}^{\infty} (-c)^k v(t-k)$$

One-step prediction of v

$$v(t) = e(t) + \sum_{k=1}^{\infty} h(k)e(t-k) \qquad (*)$$

Suppose $\{v(k)\}_{k=-\infty}^{t-1}$ has been observed. Then $\{e(k)\}_{k=-\infty}^{t-1}$ can be recovered by

$$e(t-k) = \sum_{\ell=0}^{\infty} \tilde{h}(\ell) v(t-k-\ell)$$

Calling the second term of $(*)$ $m(t-1)$,

$$m(t-1) = \sum_{k=1}^{\infty} h(k) \sum_{\ell=0}^{\infty} \tilde{h}(\ell) v(t-k-\ell)$$

and

$$v(t) = e(t) + m(t-1)$$

Define $\hat{v}(t|t-1)$ as the predicted value of $v(t)$ given all measurements through time $t-1$, and assuming $E[e(t)] = 0 \;\forall\, t$,

$$\hat{v}(t|t-1) = m(t-1) = E_{t|t-1}[v(t)] = \sum_{k=1}^{\infty} h(k) e(t-k)$$

Writing this another way,

$$\hat{v}(t|t-1) = [H(q) - 1]e(t) = [H(q)-1]\left\{\frac{v(t)}{H(q)}\right\}$$

$$= [1 - H^{-1}(q)]v(t)$$

$$= -\sum_{k=1}^{\infty} \tilde{h}(k) v(t-k)$$

Returning to our moving average example,

$$H^{-1}(z) = \sum_{k=0}^{\infty} (-c)^k z^{-k}$$

so

$$\hat{v}(t|t-1) = -\sum_{k=1}^{\infty} (-c)^k v(t-k)$$

Suppose now that a deterministic (or measured) input $u(t)$ is known, along with the output $y(t)$.

$$y(s) = G(q)u(s) + H(q)e(s)$$

Since

$$v(s) = H(q)e(s)$$

then

$$v(s) = y(s) - G(q)u(s)$$

If measurements of $u(s), y(s)$ are known for $s \le t-1$, then $v(s)$ is also known. The next output $y(t)$ is to be predicted.

Taking $E_{t|t-1}[\ ]$,

$$\begin{aligned} \hat{y}(t|t-1) &= G(q)u(t) + \hat{v}(t|t-1) \\ &= G(q)u(t) + [1 - H^{-1}(q)]v(t) \\ &= G(q)u(t) + [1 - H^{-1}(q)][y(t) - G(q)u(t)] \end{aligned}$$

Simplifying

$$\boxed{\hat{y}(t|t-1) = H^{-1}(q)G(q)u(t) + [1 - H^{-1}(q)]y(t)}$$

or, multiplying by $H(q)$:

$$\boxed{H(q)\hat{y}(t|t-1) = G(q)u(t) + [H(q) - 1]y(t)}$$

If

$$\frac{G(z)}{H(z)} = \sum_{k=1}^{\infty} \ell(k) z^{-k}$$

then

$$\hat{y}(t|t-1) = \sum_{k=1}^{\infty} \ell(k) u(t-k) - \sum_{k=1}^{\infty} \tilde{h}(k) y(t-k)$$

Unknown Initial Condition

Typically, measurements are only made from an initial time, say $t=0$, rather than for all time. The equations assume all past values are available.

A simple (and suboptimal) solution is to replace all unknown values by zero:

$$\hat{y}(t|t-1) \cong \sum_{k=1}^{t} \ell(k) u(t-k) - \sum_{k=1}^{t} -\tilde{h}(k) y(t-k)$$

There is an error between this & the optimal predictor. If G/H and H^{-1} are stable, the error decays asymptotically to zero as t increases.

The optimal solution is the Kalman filter..

Observers

Suppose there is no explicit noise model, ~~beg~~ but

$$y(t) = G(q)u(t) \qquad (**)$$

In this case, there are multiple choices for the observer

Suppose, for example, that a stable filter

$$W(q) = 1 + \sum_{\ell=k}^{\infty} w_\ell q^{-\ell}$$

is chosen. Apply in to both sides of $(**)$

$$W(q)y(t) = W(q)G(q)u(t)$$

or

$$y(t) = [1 - W(q)]y(t) + W(q)G(q)u(t)$$

We can use this to produce a "guess" of the value of y at time t:

$$\hat{y}(t|t-k) = [1 - W(q)]y(t) + W(q)G(q)u(t)$$

$W(q)$ is a free parameter to be chosen from the set of stable systems $\left(\sum_{l=k}^{\infty} |w_l| < \infty \right)$, and all such $W(q)$ produce an <u>observer</u>.

There are trade-offs in this choice. For example,

1. $W(q)$ with rapidly decaying impulse responses yield observers with errors caused by unknown inital conditions that quickly decay.
2. $W(q)$ can be chosen to use more information from past measurements in order to attenuate measurement noise effects.

The Kalman filter is a way to choose $W(q)$ in a manner that optimizes this trade-off to yield minimum error variance.