Continuous-time Least Squares Estimation

Data generated by

$$y(t) = \theta^{*T}\varphi(t) \qquad t \in \Re$$

(See also Astrom pp 71 ff)

Estimator is

$$\frac{d}{dt}\hat{\theta}(t) = -P(t)\varphi(t)e_1(t)$$
$$e_1(t) = \hat{y}(t) - y(t) = \hat{\theta}(t)^T\varphi(t) - y(t)$$
$$\frac{d}{dt}P(t) = -P(t)\varphi(t)\varphi(t)^T P(t)$$

Remarks

- Can be shown to minimise

$$J(\theta) = \int_0^t \left(\theta^T\varphi(t) - y(t)\right)^2 dt$$

- Can also be shown to have some nice stability properties.

Auxiliary result

For non-singular P(t), $\dot{P}^{-1} = -P^{-1}\dot{P}P^{-1}$

Proof: ???

Stability properties of the least squares estimator

Define $\tilde{q}(t) \overset{\Delta}{=} \hat{q}(t) - q^{*}$, then $e_1(t) = \tilde{q}(t)^T j(t)$.

Consider the quadratic form $V(t) = \tilde{q}^T(t) P(t)^{-1} \tilde{q}(t)$.

Its derivative is

$$
\begin{aligned}
\frac{d}{dt}V(t) &= 2\tilde{q}^T(t)P^{-1}(t)\dot{\tilde{q}}(t) + \tilde{q}^T(t)\dot{P}^{-1}(t)\tilde{q}(t) \\
&= 2\tilde{q}^T P^{-1}(-Pje_1) + \tilde{q}^T\left(-P^{-1}\left[-Pjj^TP\right]P^{-1}\right)\tilde{q} \\
&= -2\tilde{q}^T je_1 + \tilde{q}^T jj^T\tilde{q} \\
&= -2e_1^2 + e_1^2 = -e_1^2
\end{aligned}
$$

Further,

$$\dot{P}^{-1} = -P^{-1}\dot{P}P^{-1} = jj^T$$

Clearly

$$P^{-1}(t) = P^{-1}(t_1) + \int_{t_1}^{t} j(t)j(t)^T dt$$

$$l_{\min}\left[P^{-1}(t)\right] \geq l_{\min}\left[P^{-1}(t_1)\right] \qquad for \quad t \geq t_1$$

Since

$$\frac{d}{dt}V(t) = -e_1^2 = -\tilde{\boldsymbol{q}}^T \boldsymbol{j}\boldsymbol{j}^T \tilde{\boldsymbol{q}}$$

we have

$$V(t) = V(t_1) - \int_{t_1}^{t} e_1^2(t)\,dt$$

$$\tilde{\boldsymbol{q}}(t)^T P(t)^{-1} \tilde{\boldsymbol{q}}(t) = \tilde{\boldsymbol{q}}(t_1)^T P(t_1)^{-1} \tilde{\boldsymbol{q}}(t_1) - \int_{t_1}^{t} e_1(\boldsymbol{t})^2\,d\boldsymbol{t}$$

Thus,

$$\tilde{\boldsymbol{q}}(t)^T P(t)^{-1} \tilde{\boldsymbol{q}}(t) \le \tilde{\boldsymbol{q}}(t_1)^T P(t_1)^{-1} \tilde{\boldsymbol{q}}(t_1) \qquad \textit{for} \quad t \ge t_1$$

$$\therefore \qquad \boldsymbol{l}_{\min}\left(P(t)^{-1}\right)\left\|\tilde{\boldsymbol{q}}(t)\right\|^2 \le \tilde{\boldsymbol{q}}(0)^T P(0)^{-1} \tilde{\boldsymbol{q}}(0)$$

But

$$\boldsymbol{l}_{\min}\left(P(t)^{-1}\right) \ge \boldsymbol{l}_{\min}\left(P(0)^{-1}\right)$$

$$\therefore \qquad \boldsymbol{l}_{\min}\left(P(0)^{-1}\right)\left\|\tilde{\boldsymbol{q}}(t)\right\|^2 \le \boldsymbol{l}_{\max}\left(P(0)^{-1}\right)\left\|\tilde{\boldsymbol{q}}(0)\right\|^2$$

or

$$\left\|\tilde{\boldsymbol{q}}(t)\right\|^2 \le K \left\|\tilde{\boldsymbol{q}}(0)\right\|^2$$

Thus in fact, the least squares estimator is Lyapunov stable.

(Note that the gradient estimator is also Lyapunov stable. The proof above is clearly more difficult as the least squares estimator is more complicated.)

- Least squares estimator has a better initial convergence rate.
- Least squares estimator results in estimates with a lower variance.
- However, least squares estimator is a decreasing gain estimator.

(Try monitoring diagonal elements of P(t) in simulation.)

Statistical Propertis of LSE (Extra materials)

Suppose e_k, $k = 1,2,...t$, are zero-mean uncorrelated random variables with variance $Ee_k^2 = s^2$, and that

$$y_k = j_k^T q + e_k$$

and $Y = [y_1, y_2, ..., y_t]^T$ and $\Phi = [j_1, j_2, ..., j_t]^T$ are given. If $\Phi^T \Phi$ is non-singular, then the LS estimation $\hat{q} = (\Phi^T \Phi)^{-1} \Phi^T y$ has the following properties:

1). $E\hat{q} = q$ ($\hat{q}$ is an unbiased estimate of q)

2). The variance matrix of $\hat{q}$

$$\mathrm{cov}\hat{q} = s^2 (\Phi^T \Phi)^{-1}$$

3). Gauss-Markov Theorem: The LSE $\hat{q}$ is the Best Linear Unbiased Estimator (BLUE) in the data $Y = [y_1, y_2, ..., y_t]^T$ in the sense that for any other "LUE" $b = BY$,

$\text{cov}\,\hat{q} \le \text{cov}\, b$

Modified Least Squares Estimator (Continuous-time)

(Decreasing gain of least squares estimator is a problem as the estimator becomes insensitive after some time.)

Consider modified estimator I

$$\frac{d}{dt}\hat{q}(t) = -P(t)j\,(t)e_1(t)$$

$$e_1(t) = \hat{y}(t) - y(t) = \hat{q}(t)^T j\,(t) - y(t)$$

$$\frac{d}{dt}\overline{P}(t) = -P(t)j\,(t)j\,(t)^T P(t)$$

$$P(t) = \overline{P}(t) + aI \qquad when \quad tr(\overline{P}) < tr(P(0))$$

$$P(t) = \overline{P}(t) \qquad when \quad tr(\overline{P}) \ge tr(P(0))$$

The addition of αI acts to prevent estimator gain from decreasing to zero.

Another way to prevent decreasing gain estimator, consider criterion

$$J(\theta) = \int_0^t e^{\alpha(t-\tau)} \left(y(\tau) - \varphi^T(\tau)\theta \right)^2 d\tau \,, \quad \alpha < 0$$

This leads to modification II

Continuous-time least squares estimation with exponential forgetting

$$\frac{d}{dt}\hat{\theta} = -P(t)\varphi(t)e_1(t)$$

$$e_1(t) = \varphi^T(t)\hat{\theta}(t) - y(t)$$

$$\frac{d}{dt}P(t) = \alpha P(t) - P(t)\varphi(t)\varphi(t)^T P(t)$$

Exercise

Consider the stability properties of this estimator by examining

$$V(t) = \tilde{\theta}(t)^T P^{-1}(t)\tilde{\theta}(t) .$$

Parallel developments also possible for discrete-time.

Recursive Least Squares (See also Astrom…pp.66)

$$\hat{q}(t) = \hat{q}(t-1) - \frac{P(t-1)j(t)}{1 + j(t)^T P(t-1)j(t)} e_1(t) \quad (5.1)$$

$$e_1(t) = \hat{q}(t-1)^T j(t) - y(t) \quad (5.2)$$

$$P(t) = P(t-1) - \frac{P(t-1)j(t)j(t)^T P(t-1)}{1 + j(t)^T P(t-1)j(t)} \quad (5.3)$$

- Typically, this is mainly viewed as the estimator minimizing the criterion

$$J(q,t) = \frac{1}{2}\sum_{j=1}^{t}\left(y(j) - j^T(j)q\right)^2$$

- Estimator can also be shown to be Lyapunov stable for data generated by

$$y(t) = j(t)^T q^*$$

Consider the quadratic form

$$V(t) = \tilde{q}(t)P(t)^{-1}\tilde{q}(t)$$

- It should be clear by now that it is not sufficient to simply have a quadratic form.

- Can the above quadratic form qualify as a positive definite function if it is possible that

$$l_{\min}\left(P(t)^{-1}\right) = 0 \qquad ?$$

This will not be so in the least-squares estimator.

Demonstration of Lyapunov Stability

From matrix inversion lemma (Astrom pp.67), it can be shown that (5.3) is equivalent to

$$P(t)^{-1} = P(t-1)^{-1} + j(t)j(t)^{T}$$

$$l_{\min}\left[P(t)^{-1}\right] \geq l_{\min}\left[P(t-1)^{-1}\right] \geq l_{\min}\left[P(0)^{-1}\right] > 0$$

Clearly, V(t) is positive definite.

Further

$$e_1(t) = j\,(t)^T \tilde{q}\,(t-1)$$

and from (5.1)

$$\tilde{q}\,(t) = \tilde{q}\,(t-1) - \frac{P(t-1)j\,(t)j\,(t)^T}{1+j\,(t)^T P(t-1)j\,(t)}\tilde{q}\,(t-1)$$

$$= \left\{P(t-1) - \frac{P(t-1)j\,(t)j\,(t)^T P(t-1)}{1+j\,(t)^T P(t-1)j\,(t)}\right\}P(t-1)^{-1}\tilde{q}\,(t-1)$$

$$= P(t)P(t-1)^{-1}\tilde{q}\,(t-1)$$

i.e.,

$$P^{-1}(t)\tilde{q}\,(t) = P(t-1)\tilde{q}\,(t-1)$$

Consider now

$$V(t) - V(t-1) = \tilde{q}\,(t)^T P(t)^{-1}\tilde{q}\,(t) - \tilde{q}\,(t-1)^T P(t-1)^{-1}\tilde{q}\,(t-1)$$

$$= \tilde{q}\,(t)^T P(t-1)^{-1}\tilde{q}\,(t-1) - \tilde{q}\,(t-1)^T P(t-1)^{-1}\tilde{q}\,(t-1)$$

$$= \left\{\frac{-P(t-1)j\,(t)e_1(t)}{1+j^T(t)P(t-1)j\,(t)}\right\}^T P(t-1)^{-1}\tilde{q}\,(t-1)$$

$$= \frac{-e_1(t)^2}{1+j^T(t)P(t-1)j\,(t)}$$

Therefore,

$$V(t) = V(0) - \sum_{j=1}^{t} \frac{e_1(j)^2}{1 + j^T(j)P(j-1)j(j)}$$

$$\tilde{q}(t)P(t)^{-1}\tilde{q}(t) \le \tilde{q}(0)P(0)^{-1}\tilde{q}(0)$$

or

$$l_{\min}\left(P(0)^{-1}\right)\left\|\tilde{q}(t)\right\|^2 \le l_{\min}\left(P(t)^{-1}\right)\left\|\tilde{q}(t)\right\|^2 \le l_{\max}\left(P(0)^{-1}\right)\left\|\tilde{q}(0)\right\|^2$$

Since

$$l_{\min}\left(P(t)^{-1}\right) \ge l_{\min}\left(P(0)^{-1}\right)$$

Thus,

$$\left\|q(t)\right\|^2 \le \frac{l_{\max}\left(P(0)^{-1}\right)}{l_{\min}\left(P(0)^{-1}\right)}\left\|\tilde{q}(0)\right\|^2$$

Thus the RLS estimator is Lyapunov stable.

(In fact, stable in the large!)

We can proceed further

$$V(t) = V(0) - \sum_{j=1}^{t} \frac{e_1(j)^2}{1 + j^T(j)P(j-1)j(j)}$$

$\therefore \quad V(t) \le V(0)$

In addition, $\quad V(t) = \tilde{q}(t)P(t)^{-1}\tilde{q}(t) \ge 0$

$$\sum_{j=1}^{t} \frac{e_1(j)^2}{1 + j^T(j)P(j-1)j(j)} = V(0) - V(t)$$

$$\Rightarrow \sum_{j=1}^{t} \frac{e_1(j)^2}{1 + j^T(j)P(j-1)j(j)} \le c_1 \qquad for\, all \quad t$$

Quite easy to see that $l_{\max}(P(j)) \le l_{\max}(P(0)) \qquad \forall j$

$$\sum_{j=1}^{t} \frac{e_1(j)^2}{1 + l_{\max}(P(0))j(j)^T j(j)} \le c_1$$

$$\Rightarrow \lim_{t\to\infty} \frac{e_1(t)}{\sqrt{1 + l_{\max}(P(0))j(t)^T j(t)}} = 0$$

Further

$$\left\|\hat{q}(t)-\hat{q}(t-1)\right\|^2 = \frac{j(t)^T P(t-1)P(t-1)j(t)e_1(t)^2}{\left(1+j(t)^T P(t-1)j(t)\right)^2}$$

$$= \frac{j(t)^T P(t-1)P(t-1)j(t)}{1+j(t)^T P(t-1)j(t)} * \frac{e_1(t)^2}{1+j(t)^T P(t-1)j(t)}$$

(Remembering that $P(t)$ is symmetric positive semi-definite.)

We already know

$$\frac{e_1(t)^2}{1+j^T(t)P(t-1)j(t)} \to 0$$

Further, since $P(t-1)$ is symmetric, positive semi-definite

$$j^T PP j \le l_{\max}(P(t-1))j^T P j \le l_{\max}(P(0))j^T P j$$

$$\frac{j^T PP j}{1+j^T P j} \le l_{\max}(P(0))\frac{j^T P j}{1+j^T P j} \le l_{\max}(P(0))$$

$$\left\|\hat{q}(t)-\hat{q}(t-1)\right\|^2 \le l_{\max}(P(0))\frac{e_1^2(t)}{1+j^T P j} \to 0$$

$$as \quad t \to \infty$$

Thus, for the RLS estimator, we have

(i) $\left\|\hat{q}(j)\right\|$ bounded;

(ii) $$\lim_{t\to\infty} \frac{e_1(t)}{\sqrt{1+l_{\max}\left(P(0)\right)j^T(t)j(t)}} = 0;\text{and}$$

(iii) $$\lim_{t\to\infty}\left\|\hat{q}(t+k)-\hat{q}(t)\right\| = 0$$ for all finite k

What does this mean if we use the RLS algorithm in the minimum prediction error adaptive controller?

Again, the basic RLS is a decreasing gain estimator.

Possible modification…

Least Squares with Covariance Modification

$$\hat{\theta}(t) = \hat{\theta}(t-1) - \frac{P(t-1)\varphi(t)}{1+\varphi(t)^T P(t-1)\varphi(t)} e_1(t)$$

$$e_1(t) = \hat{\theta}(t-1)\varphi(t) - y(t)$$

$$\overline{P}(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi(t)^T P(t-1)}{1+\varphi(t)^T P(t-1)\varphi(t)}$$

$$P(t) = \overline{P}(t) + \alpha I \qquad when \qquad tr(P(t)) < k_{\max}$$

$$P(t) = \overline{P}(t) \qquad otherwise$$

Again, the addition of αI term prevents estimator gains from becoming too low.

Modification II: RLS with exponential forgetting

Another modification is to consider an estimator minimising

$$J(q,t) = \frac{1}{2}\sum_{j=1}^{t} l^{t-j}\left(y(j) - j^{T}(j)q\right)^2$$

where $0 < l \leq 1$.

- Obviously, it weights recent data more heavily.
- This criterion leads to

$$\hat{q}(t) = \hat{q}(t-1) - \frac{P(t-1)j(t)e_1(t)}{l + j(t)^T P(t-1)j(t)}$$

$$e_1(t) = j(t)^T \hat{q}(t-1) - y(t)$$

$$P(t) = \frac{1}{l}P(t-1) - \frac{1}{l}\frac{P(t-1)j(t)j(t)^T P(t-1)}{l + j(t)^T P(t-1)j(t)}$$

Robustness Considerations

- To describe behaviour of algorithms in non-ideal cases.
- To give ideas for more robust new algorithms.

Behavior of adaptive systems is quite complex due to their nonlinear character.

Very unlike linear time-invariant systems where most problems can be answered in detail.

We have already looked at stability analysis of some adaptive controllers. We will look at

- convergence of parameter estimates
- consequence of violating assumptions in stability analysis explored

Read again Astrom …Section 6.2.

Ideas should already be clear.

Effects of disturbances

Recall that the stability results were developed for plant described by

$$A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t)$$

It is likely that in real-life there will be disturbances (assume $v_1(t)$ bounded)

$$A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t) + v_1(t)$$

The corresponding estimation model will now be of the form

$$y(t) = \varphi(t-1)^T \theta^* + v(t) \qquad (6.1)$$

where $v(t)$ is bounded since $v_1(t)$ is bounded.

This disturbance term can cause us problems.

We need to modify our parameter estimation to handle this.

Robust Parameter Estimation

Gradient Algorithm with Dead-Zone

$$\hat{q}(t) = \hat{q}(t-1) + \frac{a(t-1)j\,(t-1)}{c + j\,(t-1)^T j\,(t-1)}\left[y(t) - j\,(t-1)^T \hat{q}(t-1)\right] \tag{6.2}$$

with $\hat{q}(0)$ specified, and any c>0, and

$$a(t-1) = \begin{cases} 1 & if \quad \left|y(t) - j\,(t-1)^T \hat{q}(t-1)\right| > 2\Delta \\ 0 & otherwise \end{cases} \tag{6.3}$$

Lemma 3.6.1: Consider the system model given in (6.1), where {v(t)} is a bounded sequence such that

$$\sup|v(t)| \le \Delta \tag{6.4}$$

Then the algorithm (6.2)—(6.3) has the following properties:

(i) $$\left\|\hat{q}(t) - q^*\right\| \le \left\|\hat{q}(t-1) - q^*\right\| \le \left\|\hat{q}(0) - q^*\right\|, \quad t \ge 1 \tag{6.5}$$

(ii) $$\lim_{N\to\infty} \sum_{t=1}^{N} \frac{a(t-1)\left\{e(t)^2 - 4\Delta^2\right\}}{c + j(t-1)^T j(t-1)} \le c_1 \tag{6.6}$$

where e(t) is the modelling error given by

$$\begin{aligned} e(t) &= y(t) - j(t-1)^T \hat{q}(t-1) \\ &= -j(t-1)^T \tilde{q}(t-1) + v(t) \end{aligned} \tag{6.7}$$

$$\tilde{q}(t) \overset{\Delta}{=} \hat{q}(t) - q^*$$

and this implies

(a) $$\lim_{t\to\infty} \frac{a(t-1)\left\{e(t)^2 - 4\Delta^2\right\}}{c + j(t-1)^T j(t-1)} = 0 \tag{6.8}$$

(b) $$\lim_{t\to\infty}\sup \left\|\hat{q}(t) - \hat{q}(t-1)\right\| \le \frac{2\Delta}{\sqrt{c}} \tag{6.9}$$

In addition, provided that $\{j(t)\}$ is a bounded sequence, we may conclude that

(c) $$\lim_{t\to\infty}\sup |e(t)| \le 2\Delta \tag{6.10}$$

Proof:

Subtracting q^* from both sides of (6.2) and using (6.1) gives

$$\tilde{q}(t) = \tilde{q}(t-1) - \frac{a(t-1)j(t-1)}{c + j(t-1)^T j(t-1)}\left[j(t-1)^T \tilde{q}(t-1) - v(t)\right]$$
$$= \tilde{q}(t-1) - \frac{a(t-1)j(t-1)}{c + j(t-1)^T j(t-1)}\left[-e(t)\right]$$

(6.11)

Then, noting that *a(t-1)=0* or 1, we have

$$\left\|\tilde{q}(t)\right\|^2 = \left\|\tilde{q}(t-1)\right\|^2 - \frac{2a(t-1)\left[e(t) - v(t)\right]e(t)}{c + j(t-1)^T j(t-1)}$$
$$+ \frac{a(t-1)^2 j(t-1)^T j(t-1)e(t)^2}{\left[c + j(t-1)^T j(t-1)\right]^2}$$
$$\le \left\|\tilde{q}(t-1)\right\|^2 + \frac{a(t-1)}{c + j(t-1)^T j(t-1)}\left[2e(t)v(t)\right]$$
$$- \frac{a(t-1)e(t)^2}{c + j(t-1)^T j(t-1)} \qquad (6.11a)$$
$$= \left\|\tilde{q}(t-1)\right\|^2 - \frac{a(t-1)e(t)}{c + j(t-1)^T j(t-1)}\left[e(t) - 2v(t)\right] \qquad (6.11b)$$

Notes that

$$**** \quad \frac{a(t-1)^2 j^T j\, e(t)^2}{\left[c + j^T j\right]^2} = \frac{a(t-1)^2 e(t)^2}{c + j^T j} \frac{j^T j}{c + j^T j}$$

From (6.11b), when

(i) $|e(t)| < 2\Delta, \quad a(t-1) = 0$

(ii) $|e(t)| > 2\Delta > 2|v(t)|, \quad a(t-1) = 1$

we have

$$\|\tilde{q}(t)\| - \|\tilde{q}(t-1)\| = 0 \quad \text{in case (i)}$$

$$\|\tilde{q}(t)\| - \|\tilde{q}(t-1)\| \leq 0 \quad \text{in case (ii)}$$

Net result is

$$\|\tilde{q}(t)\| \leq \|\tilde{q}(t-1)\|$$

- In case (i), when estimation error is of the magnitude of the noise, stop update;
- Only when estimation error is larger than magnitude of noise, as in case (ii), updating proceed.

This proves (i).

Proceeding from (6.11a)

$$\begin{aligned}
\left\|\tilde{q}(t)\right\|^2 &\le \left\|\tilde{q}(t-1)\right\|^2 + \frac{a(t-1)}{c+j(t-1)^T j(t-1)}\left[2e(t)v(t)\right] \\
&\quad - \frac{a(t-1)e(t)^2}{c+j(t-1)^T j(t-1)} \\
&\le \left\|\tilde{q}(t-1)\right\|^2 + \frac{a(t-1)}{c+j(t-1)^T j(t-1)}\left[\frac{e^2(t)}{2} + 2v(t)^2\right] \\
&\quad - \frac{a(t-1)e(t)^2}{c+j(t-1)^T j(t-1)} \qquad (6.12)
\end{aligned}$$

(since $2ab \le ka^2 + b^2/k$, for any k>0)

$$\begin{aligned}
&\le \left\|\tilde{q}(t-1)\right\|^2 - \frac{1}{2}\frac{a(t-1)e(t)^2}{c+j(t-1)^T j(t-1)} \\
&\quad + \frac{2a(t-1)\Delta^2}{c+j(t-1)^T j(t-1)} \qquad (6.12a)
\end{aligned}$$

Thus, $\dfrac{1}{2}\dfrac{a(t-1)\left\{e(t)^2 - 4\Delta^2\right\}}{c+j(t-1)^T j(t-1)} \le \left\|\tilde{q}(t-1)\right\|^2 - \left\|\tilde{q}(t)\right\|^2$

Summing both sides and using (i) proves (ii).

Since $a(t-1)\left\{e(t)^2 - 4\Delta^2\right\} \geq 0$ (because of the choice of a(t-1)), the result (a) follows from (ii).

To establish (b), we note that from (6.2) and (6.7) that

$$\hat{q}(t) - \hat{q}(t-1) = \frac{-a(t-1)j\,(t-1)}{c + j\,(t-1)^T j\,(t-1)} e(t)$$

Hence

$$\left\| \hat{q}(t) - \hat{q}(t-1) \right\|^2 = \frac{a(t-1)j\,(t-1)^T j\,(t-1)}{\left[c + j\,(t-1)^T j\,(t-1)\right]^2} e(t)^2 \qquad (6.13)$$

From (6.8)

$$\lim_{t\to\infty} \sup \frac{a(t-1)e(t)^2}{c + j\,(t-1)^T j\,(t-1)} = \lim_{t\to\infty} \frac{4a(t-1)\Delta^2}{c + j^T j} \leq \frac{4\Delta^2}{c}$$

(6.14)

For $0 < \dfrac{j\,(t-1)^T j\,(t-1)}{c + j\,(t-1)^T j\,(t-1)} < 1$, we have

$$\lim_{t\to\infty} \sup \frac{a(t-1) j\,(t-1)^T j\,(t-1) e(t)^2}{\left[c + j\,(t-1)^T j\,(t-1)\right]^2} \le \frac{4\Delta^2}{c}$$

(Note that $\dfrac{j^T j}{\left[c + j^T j\right]} * \dfrac{ae^2}{\left[c + j^T j\right]}$)

and from (6.13), and the above, (b) follows.

If {φ(t)} is a bounded sequence, then (c) follows from (6.8), i.e. result (a), or (6.14).

The above result can be used in the context of adaptive control in the presence of bounded disturbances.

See also Narendra….pp. 323

Astrom… pp226 ff

for alternative viewpoint.

Least Squares with Dead Zone

$$\hat{q}(t) = \hat{q}(t-1) + \frac{a(t-1)P(t-1)j\,(t-1)}{1 + a(t-1)j\,(t-1)^T P(t-1)j\,(t-1)} e(t)$$

$$e(t) = y(t) - j\,(t-1)^T \hat{q}(t-1)$$

$$P(t) = P(t-1) + \frac{a(t-1)P(t-1)j\,(t-1)j\,(t-1)^T P(t-1)}{1 + a(t-1)j\,(t-1)^T P(t-1)j\,(t-1)}$$

with $\hat{q}(0)$ and $P(0) = P_0 > 0$ specified,

$$a(t-1) = \begin{cases} 1 \quad if & \dfrac{e(t)^2}{1 + j\,(t-1)^T P(t-1)j\,(t-1)} > \Delta^2 > 0 \\ 0 \quad otherwise & \end{cases}$$

Constrained Parameter Estimation

For the estimation model

$$y(t) = j\,(t-1)^T q^* + v(t) \qquad (6.15)$$

with the standard gradient estimator

$$\hat{q}(t) = \hat{q}(t-1) + \frac{j\,(t-1)}{c + j\,(t-1)^T j\,(t-1)} e(t)$$

$$e(t) = y(t) - j\,(t-1)^T \hat{q}(t-1)$$

Slight modification of the preceding analysis will show that

$$\left\|\tilde{q}(t)\right\|^2 - \left\|\tilde{q}(t-1)\right\|^2 \le -\frac{e(t)\left[e(t) - 2v(t)\right]}{c + j\,(t-1)^T j\,(t-1)}$$

(not less than or equal to 0)

i.e. as a result of the disturbance, parameter errors cannot be guaranteed to be non-increasing.

The dead-zone is one way to handle that, and it leads to

$$\left\|\tilde{q}(t)\right\|^2 - \left\|\tilde{q}(t-1)\right\|^2 \le -\frac{a(t-1)e(t)\left[e(t)-2v(t)\right]}{c + j(t-1)^T j(t-1)}$$
$$\le 0$$

(because of the choice of a(t-1).)

- However, that needs knowledge on bound on {v(t)} disturbance (equivalent disturbance in estimation model) sequence, i.e.,

  $$\sup_t \left|v(t)\right| \le \Delta$$

  and Δ known.

Suppose, instead of knowledge on Δ, we know instead bounds on the individual elements of q^*:

$$q^*_{i_{\min}} \le q^*_i \le q^*_{i_{\max}}$$

Thus, we do not know q^*, but we have some idea of the possible range of values.

Consider the plant model, $q^* \in \Re^n$

$$y(t) = j\,(t-1)^T q^* + v(t) \tag{6.16}$$

<u>Constrained estimator:</u>

$$\hat{q}_1(t) = \hat{q}\ (t-1) + \frac{j\,(t-1)}{c + j\,(t-1)^T j\,(t-1)} e(t)$$

$$e(t) = y(t) - j\,(t-1)^T \hat{q}(t-1)$$

for $1 \le i \le n$

$$\hat{q}_i(t) = \begin{cases} \hat{q}_{1i}(t) & if \quad \hat{q}_{i_{\min}}(t) \le \hat{q}_{1i}(t) \le \hat{q}_{i_{\max}}(t) \\ \hat{q}_{i_{\min}} & if \quad \hat{q}_{1i}(t) < \hat{q}_{i_{\min}} \\ \hat{q}_{i_{\max}} & if \quad \hat{q}_{1i}(t) > \hat{q}_{i_{\max}} \end{cases}$$

- Key idea: if the update $\hat{q}_{1i}(t)$ falls inside the known region, then use that value. Otherwise, use the value in that subspace (of dimension one) that does not increase the parameter error in that subspace.

- Derivation of Properties

  (i) By the constraint mechanism, by design, $\left\|\tilde{q}(t)\right\|$ is bounded, where $\tilde{q}(t) \overset{\Delta}{=} \hat{q}(t) - q^*$.

  (ii) Observe next that

$$\left\|\tilde{q}(t)\right\|^2 = \sum_{i=1}^{n} \left|\hat{q}_i(t) - q_i^*\right|^2$$

The constraint procedure ensures

$$\left|\hat{q}_i(t) - q_i^*\right| \le \left|\hat{q}_{1i}(t) - q_i^*\right|$$

$$\therefore \quad \left\|\tilde{q}(t)\right\|^2 \le \left\|\tilde{q}_1(t)\right\|^2 \qquad (6.17)$$

From (6.12), it should be clear that

$$\|\tilde{q}_1(t)\|^2 \le \|\tilde{q}(t-1)\|^2 - \frac{1}{c + j(t-1)^T j(t-1)}\left[\frac{e(t)^2}{2} - 2v(t)^2\right] \qquad (6.18)$$

Noting that $\|\tilde{q}(t)\|$ is bounded, and using (6.17), we have

$$\frac{1}{2}\frac{\{e(t)^2 - 4\Delta^2\}}{c + j(t-1)^T j(t-1)} \le \|\tilde{q}(t-1)\|^2 - \|\tilde{q}(t)\|^2$$

so that

$$\frac{1}{2}\sum_{t=1}^{N}\frac{\{e(t)^2 - 4\Delta^2\}}{c + j(t-1)^T j(t-1)} \le \sum_{t=1}^{N}\left\{\|\tilde{q}(t-1)\|^2 - \|\tilde{q}(t)\|^2\right\}$$
$$= \|\tilde{q}(0)\|^2 - \|\tilde{q}(N)\|^2$$

i.e.

$$\lim_{N\to\infty}\sum_{t=1}^{N}\frac{\{e(t)^2 - 4\Delta^2\}}{c + j(t-1)^T j(t-1)} \le c_1$$

Thus, it has some properties similar to the dead-zone method.

- Knowledge of a bounded region for q^* can also be used in a similar way in continuous-time estimation.

Comparison of dead-zone and constrained parameter estimation

- Prior knowledge: dead-zone requires knowledge of Δ; constrained method requires knowledge of $\left[q^*_{i_{\min}}, q^*_{i_{\max}}\right]$, i=1,2,…,n.
- Both methods ensure boundedness of $\hat{q}(t)$
- Dead-zone ensure $\|\tilde{q}(t)\|^2 \le \|\tilde{q}(t-1)\|^2$. Constrained method only ensures boundedness.
- For dead-zone method, once Δ is fixed in the rule for a(t-1), update stops when e(t) satisfying $|e(t)| \le 2\Delta$. Thus accuracy limited even when disturbance dies away. In constrained method, when v(t)=0, estimation "improves".

Matrix Inversion Lemma

Let A, C and $C^{-1}+DA^{-1}B$ be nonsingular square matrices. Then,

$$(A+BCD)^{-1}=A^{-1}-A^{-1}B(C^{-1}+DA^{-1}B)^{-1}DA^{-1}$$

Proof: Multiply by (A+BCD) and verify the identity.