1. Static System Identifications
   1. Linear Static Systems
      1. Linear Regression

* Definition: linear regression
* (5.1) linear regression model
* is an observed output, regressor, an unknown parameter
* Why regression? The opposite of regression is (advancement, development, evolution..)
* Ex. 5.2: Moving object

Model: the distance

Observations of the distance

* + 1. Least Squares Estimation
* Def: the prediction error(or residuals)
* Problem statement:

find the least square error estimator of

* *In stochastic control minimum variance error estimator*:

So in order to get the best MV we need to have the conditional PDF. In the least square estimator, we do not need the conditional estimator.

* Solution

Let (5.3) in the matrix form as

Hence

The gradient of is zero iff , such that

We call as LSE as

In the case of the outliers: the weighted LSE, the weight =

* Ex.5.3

Ex. 5.3 / 5.4 : Moving object

Model: the distance

Observations of the distance

1)The mean of the samples of prediction error(residual)

The samples of the prediction error

The mean of the samples of the prediction error is unbiased.

2)The sample variance of the prediction error

%% Kim’s comment : *Statistic: see the sample random variable*.

In statistics,

1. The sample mean is an estimator of the mean:

Since

* , which implies the mean of the samples are equivalent to the real value.

1. The mean of the samples of the variance is an estimator of the variance

where

1. So, The sample mean of the R.V. as N is large may be called as the mean of the R.V.

And the sample variance of the R.V. may be called as the variance of the R.V.

The variance is the same as the mean square error.

%%%%%%%%%%%

* The outlier: See the textbook p.65 (the weighted least square estimator:

-the cost function

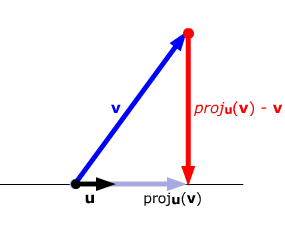
-the solution:

* + 1. Interpretation of LSE

1. Orthogonal property between error and estimator

* Ex. 5.5 Orthogonal projection

<https://www.math.uh.edu/~jiwenhe/math2331/lectures/sec6_3.pdf>



=the orthogonal projection vector of on vector

* Inner product : In Euclidean space

Then the inner product of is

* Orthogonal basis : , then is orthogonal to

Let pick up an orthogonal basis as

Then

* The orthogonality between the prediction error and the estimator

The inner product of

They are orthogonal, hence,

* Definition:

-Projection matrix:

-Orthogonal projection matrix:

-Ex. Let then and hence P is an orthogonal projection matrix. Since

Which implies the estimator is the projection of on the space of

y

* Cross-correlation of least square estimator

Implies

* + 1. Bias of LSE
* Definition: a biased or unbiased estimator

Let an estimator of . The bias is defined as

The unbiased estimator if , otherwise the biased estimator.

* **Ex.5.6: This estimator is not the LSE but an unbiased estimator)**

Define an estimator is

Then

Hence

Hence

- If then the estimator is unbiased.

- Or if is large, the estimator converges to be unbiased

- Or if is large, the estimator converges to be unbiased

* **Ex.5.7(Check) The estimator is not LSE and is not unbiased**
* The LSE is unbiased in the system

**Hence the LSE is unbiased** if

1. is independent and
2. , where

Or if is deterministic, i.e.,

Hence, which implies the estimator is always unbiased.

* **Ex.5.8 The estimator is LSE and is not unbiased**
* **Ex. 5.9 Errors – in- Variables LSE is biased**

If and are correlated.

* + 1. Accuracy (Variance) of LSE
* The covariance of

If is a white noise with constant variance , then .

(5.31) is

**Since is unknown, the variance of the prediction error**

The prediction error sequence

Thus in the variance of the estimator in (5.32) will be replaced by

%%% kim’s comment

The is found, the accuracy should be checked. One method is the standard deviation which is the square root of the variance. If a R.V. is a Gaussian, then

Hence, if is not reasonable, it is not accurately estimated. %%%

* and Ex. 5.3

Hence the only the second can be accurately estimated.

%%%%% Kim’s comment on the comparison between

Keesman page 77, (5.34) will be deleted.. We may consider later , however

1. Least Square estimator
2. The
3. Hence from (a), the variance of LSE is larger than MS.
4. If , which implies is deterministic, then the estimator of LSE is the same as that of MV.

%%%%

%% Kim’s comment : do not confuse followings:

1. The notation :

* : the estimator . : the mean
* ,

1. If the regressor in is deterministic, The LSE is the MS estimator.

But not vice versa, since is a Random variable,

1. The minimum MSE is .

Hence we should know about

If

It is sufficient to know only the initial pdf of due to Markov property

1. The adaptive algorithm

If the system matrix is time varying, so indexed as

That means if the system is operate in a different environment from the previous,

i.e., the disturbance is happened so that the , we may have to change the algorithm to estimate the state.

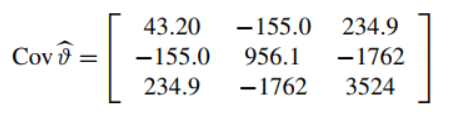
%%%%%%%

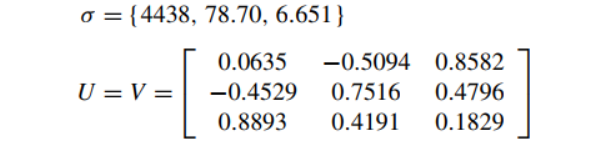
* + 1. Identifiability – application of the singular value decomposition
* Ex. 5.17

Let the covariance of then the singular value decomposition of C is

Hence if is the same direction to corresponding to the smallest sigma

its variance is smallest, which implies the most accurately estimate.

For example Ex.5.3, the covariance of

 And its sigma value and Left/Right Sigma value matrix is

Then the third column of direction is the most accurately estimated, i.e.,

%%% Kim’s comment

In stochastic, one of application of svd is Ch.4 problem 6. In there, the most (or the least) observable state or state combination is determined by SVD. See it

%%%

%%% Kim’s comment

If then the

Where

%%%

* 1. Non-linear Static Systems
     1. Non-linear Regression

where

measurements

the known model

the unknown parameters to be estimated

the prediction error ( the measurement noise)

* In the linear case,
  + 1. Non-linear Least square estimation
* The cost function (the sum of squares of the prediction errors)
* The derivative of w.r.t unknown variables (to be estimated)
* The sensitivity function

the number of the measurements

the number of the estimated aprameters

where

the first measurement with

the Nth measurement with

* In Linear case

which is independent of .

%%% Kim’s comment

Since in (a), in general, is dependent of . In concept,

First estimation

Second estimation

y(1)

y(3)

y(2)

x(1) x(2) x(3)

Hence for each , the estimation is evaluated at times with different recursively.

* In linear case

which is independent of



%%%%%%

* Apply Least square as
* Ex. 5.21 : Nitrification experiment

1. Modelling

Input:

-Nitrogen load

Output:

-The maximal oxygen demand rate

Unknown

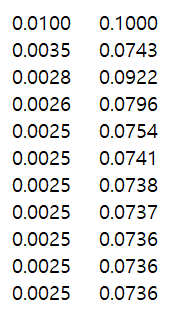
-Death rate of the nitrifying biomass :

-Maximal growth rate of the nitrifying biomass :

1. Formulation

* Define

1. Calculate the sensitivity
2. Results



Estimated (b , umax)

