1. Least Squares, The Orthogonal projection Lemma, and the Discrete-Time Kalman Filter

In chapter 3, the process and measurement are Gaussian random processes, so that the “minimum mean variance estimator” is the conditional expectation. The get the estimator, first, the conditional probability is derived based on the Gaussian pdf.

Now, in this chapter, we may design an estimator in the different way, such that the conditional pdf may not be derived. Hence the process and measurement may not be **Gaussian processes!!**

* 1. Least Squares
* The problem :
* A sequential measurements , which is

Minimize the estimator of to satisfy the cost function as

%%% Kim’s comment: least square estimator

Example: Find the straight line such that the error is minimized

%%%

* Least Square - Generalization:

The least square cost of function is

To find to minimize the cost function, we need the derivative of the function with respect to the vector

%%% Vector Calculus

-example\_1

And

-example\_2

%%%

* The vector derivative of the cost function

First

Then take the derivative at both sides

Since

Imply to

Leads to

Or using “Calculus of Variation”

🡪

* 🡪

%%% Kim’s comment

The linear equation :

1. The equation has the unique solution
2. The equation has infinite many solution
3. The equation has non solution 🡪 pseudo inverse (satisfying the min least square errors)

* If exists,
* **Ex. 4.5**

1. Condition:

The position of a vehicle are measured as

1. Problem : estimate initial position, velocity, acceleration, using LSE

-original data

-measured data

The

1. Method\_1:

Model: , find the estimator of

Model: find the estimator of

1. Results.

-.The estimators of Model\_1 are good. The estimators of Model\_2 are bad.

-.If the measurement of noise is larger(the variance), the estimator is not good in both case, Model\_1 and Model\_2.

* Def. 4.1 A matrix, , is called an orthonormal matrix if

%%% Kim’s

Ex.

1. Orthogonal : each column vectors of are orthogonal,
2. Orthonormal:
3. Orthonormal matrix is a square matrix, #of rows = #of columns

%%%

* **Theorem 4.2 (singular value decomposition)** Let be a real matrix. There exists an orthonormal matrix and an orthonormal matrix and an real matrix,such that

is of the form

where is a real and

The real **scalars**  are called the singular values of

This is better to understand as a definition of the singular value decomposition

%%% Kim’s comment :eigen value / vector

For a square matrix ,

where

Hence :

1. Eigenvalue : the matrix should be a square, but SVD is not needed.
2. Eigenvalue is different from the singular value even if the given matrix is a square.
3. Some square matrices (i.e. positive definite) they are closely related.

i.e., Given matrix A, the squre root of eigen values of is the sigma value

1. Example

Example

* Interpretation of SVD

Now

* For input , the only will be effective ,
* For output the only ( with a scale will be measured.
* For the biggest the input is the biggest output in the direction of
* For the smallest the input is the smallest output in the direction of , i.e., it may be the least observable in the output
  1. The orthogonal projection lemma
* Gram-Schmidt ortho-normalization

1. Problem: In a n-dimensional vector space ,

Given a basis

are independent but not orthogonal. Find a basis which is orthogonal to

each other.

1. Gram-Schmidt procedure:
2. For any basis, we have an orthogonal basis based on the given basis.

* Hilbert space

1. If is a Hilbert space and if then can be described by a set of basis vector, , : the number of basis may be infinite.
2. Euclidean space: the real space(not abstract space), the number of basis is finite.

* is a real inner product space, and then

1. iff
2. (ax + by, z) = a(x,z) + b(y,z)

* Norm: , measure, a notion of distance / length in Hilbert space
* What is Hilbert Space?

1. Euclidean Space is a subspace of Hilbert Space.
2. Special case: Let a continuous function space on [0, T]

Then

The basis =

where the inner product is defined as:

1. Inner product in probability space:

Two random variables, , the expectation of two RVs is an inner product.

i.e.,

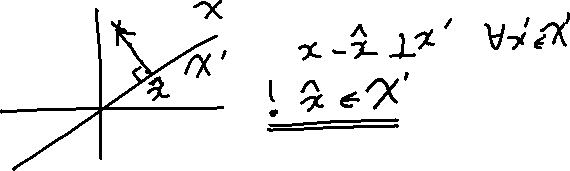
In this case iff

So the RVs ,

* **Lemma 4.7 (Orthogonal Projection Lemma**) : Let be a Hilbert space with and let be a subspace of . Then there exists a unique vector such that

Iff

* **Implication of the Lemma: the best possible estimate of a function, or a vector, is the one that is orthogonal to its estimation error.**



* 1. Extensions of Least Squares Theory(skip)
  2. Non-linear Least Squares:
* Non-linear : Newton-Gauss iteration

1. Problem

Find the Least Square Estimator for with the measurement

1. Linearization:

At the initial guess, ,

1. The cost function with the linearized approximation

Here

Then the initial estimator is

Where

Now the second estimator is

Then So that

Where

Now proceed until the iteration up to

1. Newton-Gauss iteration

%%% Kim’s comment

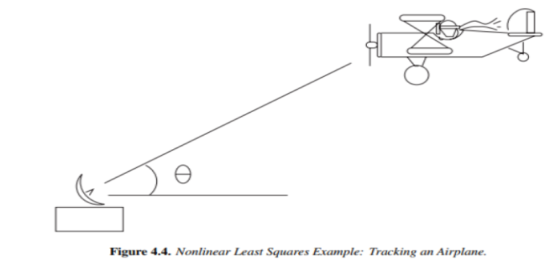
* The pseudo inverse
* The matrix is not constant since for is changed at very

It is called as Jacobian.

* As the measurement

Also is changed for

* Example. 4.9 (Tracking an airplane)



1. Problem :

* An airplane travels at a constant speed in the vertical plane
* The measurement is only the sight(angle) of a radar station
* To estimate the airplane’s position

1. Modeling

* System dynamics
* Measurements
  1. Deriving the Kalman Filter via the Orthogonal Projection Lemma
* Problem :

Where the measurement noise are assumed to be zero-mean, uncorrelated but **not necessarily Gaussian.**

1. Measurement space: Define to be a vector space , a subspace of the measurement space, as
2. The cost function:

* **Orthogonal** Example
* **projection**
* We know the best estimator should satisfy the orthogonal projection lemma

Hence the best estimator (in this case least square sense), satisfy

1. The first step
2. Apply orthogonal Projection: it is necessary

Where is any arbitrary linear function of , which is in the measurement space.

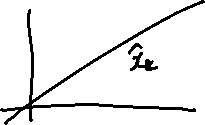
1. The orthogonal basis representation of

Let define an orthogonal basis of as

Which is

From the orthogonal projection lemma, is optimal iff

Or



which implies

* *Basis representation* :

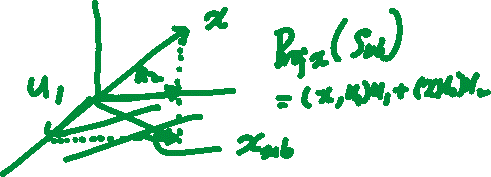
An orthonormal basis = {(1,1),(1,-1)}

Let find x = a(1,1) + b(1,-1)

* Solution:
* Projection of on the subspace

Let Define a subspace , is a subspace of , whose a basis . Then the projection of on the subspace is

Let , then



* *Hilbert Space introduced.(4.29.1)*

Since , A vector presentation of is

1. Recursive way concerning to the system
2. The innovation process:

What is the which is a basis of the measurement space of

We will pick as

which is orthogonal to

Proof: Applying the orthogonal projection lemma, it is

Multiply by

Multiply by

Therefore



is orthogonal to which is a valid choice

1. Kalman Gain



In (4.33)

We need the proportional gain. To begin with

From (4.32.1)

Define And Then

.

where

and

We then get

1. The error covariance

* Implications:

The error is orthogonal to the measurements

Since The error is orthogonal to the

1. Define the innovation as

Then

And

* Summary of Ch.4

Using Orthogonal projection Lemma, find the Kalman gain

1. The cost function:
2. Using the Lemma
3. Guess , the orthonormal basis of the measurement space
4. **Find , which is the Kalman Gain**

Where

1. For calculation of , we need an iterative method. As

Given initial conditions as

+ Initial state conditions (mean, variance):

+ Initial system parameters:

+ Initial noise parameters:

Step\_1: Prediction Step

+ mean:

+ variance:

Step\_2: Estimation (After measurement)

+ mean**:**

**+** variance**:**