1. Random Variables and Stochastic Process
2. Conditional Expectations and Discrete-Time Kalman Filtering
3. Least Squares, The Orthogonal projection Lemma, and the Discrete-Time Kalman Filter
   1. Least Squares

* The problem :
* A sequential measurements , which is

Minimize the estimator of to satisfy the cost function as

1. Generalization:

The least square cost of function is

1. Variation calculus

🡪

1. The normal equation : decomposition of vector space of

To satisfy (4.2)

* Range space (column space)

Let a matrix. The range space of is the span of the vectors , i.e.,

Range space = span of ,

* a linear combination of any vector of the form
* the set of all possible linear combinations of is called the range space of
* Pseudoinverse of : If we select as

It may be an pseudoinverse matrix of .

* Inverse matrix

1. The solution of least squares

Hence

* Normal equation in least squares.

A pseudoinverse may be

Then from

The normal equation: the solution of least squares is

which is called the normal equation .

* Def. 4.1 A matrix, , is called an orthonormal matrix if
* It is a square matrix
* The ith and jth columns of , then,
* Def 4.2 Let be a real matrix. There exists an orthonormal matrix 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 and (4.6) is the singular value decomposition of .

Hence

* The singular values of can be ordered as

And the remaining singular values are precisely zero,

Then

* Range (H) = span
* Kernel
* Kernel
* Range
* The columns of U: left singular vector,

The columns of V: right singular vector.

* The null space (kernel)

The set of all vectors such that

* The pseudoinverse of Singular value decomposition.

Then the pseudoinverse of is

* Ex.4.5 least square solution is dependent of the number of parameters and the noise intensity
  1. The orthogonal projection lemma
* Gram-Schmidt ortho-normalization

1. Problem: In a n-dimensional vector space , any independent set of n vectors turn into an orthonormal basis set. Let be a set of n independent vectors in
2. Gram-Schmidt procedure

2.1) Pick any first vector to be normalized to

where is the norm of , i.e.,

2.2) The next basis vector:

Calculate

where is the inner product The is orthogonal to since

Hence to normalize

2.3) the third basis vector

Calculate

The is orthogonal to , since

Hence to normalize

2.4) next basis vector until

* 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
* 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

Proof :

(🡨) Suppose . Then if

For equality, is necessary, which implies that that minimizes

is

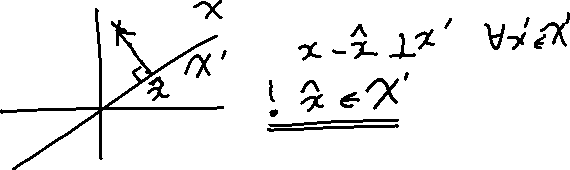
(🡪) Suppose that , and there exists an such that

Now, ,

Choose Then

But this contradicts the assumption that minimizes . QED

* **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
* Weighted Least Squares

As the previous is

The cost function is

The Weighted Least squares estimator is

Let use the Cholesky factorization as , The weighted cost function is

The necessary condition for the optimality is

* 1. **(Skip)**Non-linear Least Squares:
  2. 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 projection

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

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

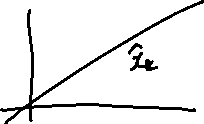
1. The orthogonal basis representation of

Let define an orthogonal basis of as

Which is

From the orthogonal projection lemma, is optimal iff

Or



which implies

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

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