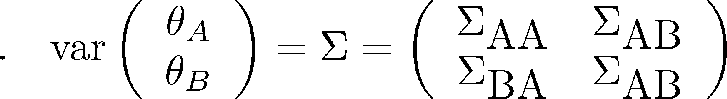
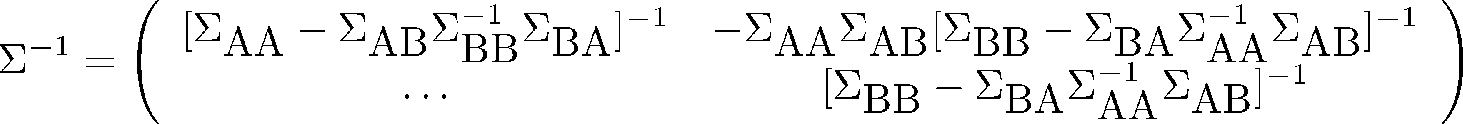
This post was published to Yifan&#039;s Blog at 2:04:30 PM 3/30/2016

Nuisance Parameters(1)

## 1 Matrix Inverse in Block Form

Assume , then it is easy to verify

 (1)

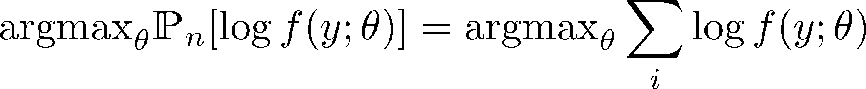
Hence, when  is considered as nuisance parameter, one can always use the  part to calculate the asymptotic variance of. Notice, there is a pseudo inverse form too.

## 2 Nuisance Parameters

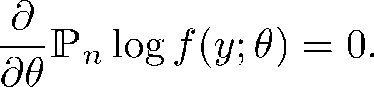
We consider a parametric model with likelihood:



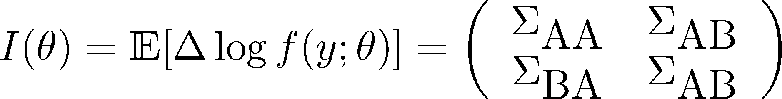
,where , and  is the pdf/mass function. We concentrate on MLE  is:



or the root of:



The fisher information used in inference for MLE is:



**Definition 1** *Orthogonal of two variables in L-2 Space*

**

Hints: In our case . Our goal now is simple:

find a way to transform dependent  to  that are Orthogonal.

It is possible to do this. (Cox and Reid, 1987) had proved  such that



and . This guides us to solve nuisance parameters’ problem by profiling them.

### 2.1 MLE Inference without nuisance parameters

Under proper conditions, .

### 2.2 MLE Inference with nuisance parameters: Profile method

Given , we could solve MLE of  by solving , which is denoted as . The profile likelihood function is defined as:



**Lemma 1** *Asyptotic distribution of Profile likelihood*

*Assume , then under proper conditions*

*(a) LRT*

*log-likelihood ratio statistic .*

*(b) Asymptotic result of *

*by* second order Taylor expansion/Jacobian and (1):



***Proof*** See (Cox and Reid, 1987), and extensions in (Fraser, Reid and Wu, 1999).

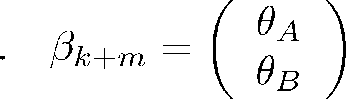
Hints:

This lemma fails when parameter space changes (e.g. with data/sample size, etc).

**2.3 MLE Inference with nuisance parameters: Conditional/Marginal**

This is a special case we could transform the problem into a ‘conditional’ version. For example, (Fellman, 1976) describes a linear regression problem:



Here we use a general setting:  needs not to be normal. if we assume , then the design matrix could be partitioned as  correspondingly. Then consider  leads to an extension of BLUE given in (Lemma 2.1 and Th 2.3) of (Fellman, 1976), which indeed only uses the pseudo inverse in block form of (projection) matrix (used in BLUE, similar to (1)).

## Ref

[1] D.R. Cox and N. Reid, “Parameter Orthogonality and Approximate Conditional Inference”, J. R. Statist. Soc. B, 47, 1, 1987.

[2] Fraser, Donald Alexander Stuart, N. Reid, and J. Wu. "A simple general formula for tail probabilities for frequentist and Bayesian inference." Biometrika 86.2 (1999): 249-264.

[3] Fellman, Johan. "On the effect of" nuisance" parameters in linear models." SankhyÄ : The Indian Journal of Statistics, Series A (1976): 197-200.