**Linear Regression Model:**

y: n by 1 vector of phenotypes/outcome/responses

X: n by p matrix of covariates/explanatory variables

β: p by 1 vector of coefficients/regression parameters

ε: n by 1 vector of residual errors

**Ordinal lease squares (OLS)**

To obtain OLS estimates, we minimize the sum of squared residuals (SSR); also called the error sum of squares (ESS) or residual sum of squares (RSS):

SSR=

To minimize this quantity, we simply take the first derivative

And set it to be zero:

When we make two assumptions: E(ε)=0; V(ε)=

Then we can derive several key properties of :

1. Unbiased: =

2.

3. is the best linear unbiased estimator (BLUE):

has the smallest variance among all unbiased estimators

that are linear functions of y (i.e. of the form , where C is p by n matrix)

based on the Gauss-Markov theorem

4. asymptotically normal (by CTL):

Prediction: with , we obtain

With X, we can obtain

Here: is an n by n matrix called the hat/projection matrix,

which projects to the column space of X

Properties of H:

1. H is idempotent: HH=H

2. I-H is also idempotent: (I-H)(I-H)=I-H

3. H(I-H)=0

4. HX=X, (I-H)X=0 (X is invariant under H)

Estimate :

**Maximum likelihood estimate (MLE)**

Assumption:

Likelihood:

Log-likelihood:

To obtain MLE, we obtain first derivatives and set them to 0:

**REML: restricted/residual maximum likelihood**

One approach to derive REML estimator is to integrate out β

To obtain REML estimate for , we can take the first derivative on the log(Lr) and we will set derivative to be zero

Therefore, .

For estimation, we simply plug in the , and obtain

Hypothesis Test: , . Here, is an r by 1 vector of transformed parameters (a.k.a. secondary parameters). is defined by L, which is called the contrast matrix. is usually a vector of zeros.

**Wald Test**

If is known

Therefore,

To test , we just need to compute the following Wald statistic that follows exactly a chi-square distribution:

If is not known, to test , we just need to compute the following Wald statistic that follows asymptotically a chi-square distribution:

**Likelihood Ratio Test**

is the maximized likelihood under unrestricted model

is the maximized likelihood under

asymptotically

**Score Test**

Define the score function: , which is an r-vector

And information matrix: , which is an r by r matrix

is the score statistics evaluated at under the null, and asymptotically

All three tests are asymptotically equivalent

In linear models, Wald statistics >= Likelihood ratio statistics >= score statistics

The likelihood ratio test often times has the best small-sample properties