MATHEMATICS

Get a Fisher information matrix for linear model with the normal distribution for measurement error?

Asked 6 years, 8 months ago Active 2 days ago Viewed 8k times

4

6

For given linear model $y = x\beta + \epsilon$, where β is a p-dimentional column vector, and ϵ is a measurement error that follows a normal distribution, a FIM is a $p \times p$ positive definite matrix.

How to find elements of the matrix?

statistics information-geometry

edited Dec 14 '13 at 9:07
Shuchang
8,827 4 19 39

asked Mar 12 '13 at 21:25
caspik
41 1 2

2 Answers

8

I'm going to assume that the variance σ^2 is known since you appear to only be considering the parameter vector β as your unknowns. If I observe a single instance (x, y) then the log-likelihood of the data is given by the density

$$\ell(\beta) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{-(y - x^T\beta)^2}{2\sigma^2}.$$

This is just the log of the Gaussian density. The Fisher information matrix is just the expected value of the negative of the Hessian matrix of $\ell(\beta)$. So, taking the gradient gives

$$S(\beta) = \nabla_\beta \frac{-(y - x^T\beta)^2}{2\sigma^2} = \nabla_\beta \left[-\frac{y^2}{2\sigma^2} + \frac{yx^T\beta}{\sigma^2} - \frac{\beta^T x x^T \beta}{2\sigma^2} \right] = \frac{yx}{\sigma^2} - \frac{xx^T\beta}{\sigma^2} = \frac{(y - x^T\beta)x}{\sigma^2}.$$

Taking another derivative, the Hessian is

$$H(\beta) = \frac{\partial}{\partial\beta^T}\frac{(y - x^T\beta)x}{\sigma^2} = \frac{\partial xy}{\partial\beta^T} - \frac{\partial xx^T\beta}{\partial\beta^T} = \frac{-xx^T}{\sigma^2},$$

so the Fisher information is

$$I(\beta) = -E_\beta H(\beta) = \frac{xx^T}{\sigma^2}.$$

Because gradients and Hessians are additive, if I observe n data items I just add the individual Fisher information matrices,

$$I(\beta) = \frac{\sum_i x_i x_i^T}{\sigma^2},$$

which, if $X^T = (x_1, x_2, \ldots, x_n)$, can be compactly written as

$$I(\beta) = X^T X/\sigma^2.$$

It is well-known that the variance of the MLE $\hat{\beta}$ in a linear model is given by $\sigma^2(X^T X)^{-1}$, and in more general settings the asymptotic variance of the MLE should be equal to the inverse of the Fisher information, so we know we've got the right answer.

edited 2 days ago
Don Thousand
7,508 3 18 39

answered Sep 3 '14 at 1:45
guy
3,248 1 20 29

0

Let γ denote the gaussian distribution of ϵ. The likelihood of the model is

$$\gamma(y - x\beta)$$

where y is your observation and β is the parameter. You can now apply the definition of the Fisher Information matrix,

$$I = \operatorname{var}\left(\nabla_\beta \log \gamma(Y - x\beta)\right).$$

answered Mar 12 '13 at 21:44
roger
2,508 10 14

Thanks for your reply. Still do not understand how the elements on main and sub diagonals will look like? – caspik Mar 12 '13 at 22:10