Multivariate linear Regression

We have r $n \times 1$ response vectors $[y_1, \ldots, y_r] = Y$
and $n \times q$ covariate matrix $X = [\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_q]$

we want to find $B_{q \times r}$ such that $Y \approx XB$.
by minimizing $\min_B \|Y - XB\|_2^2 = \sum_{\ell=1}^{r} \sum_{i=1}^{n} (y_{i\ell} - x_i' b_\ell)^2$

$$\|Y\|_2^2 = \sum_{\ell=1}^{r} \sum_{i=1}^{n} y_{i\ell}^2$$

residual sum of squares

Differentiating with respect to B yields

$$(X'X)\hat{B} = X'Y \underset{(\text{if } X \text{ is full rank})}{\Rightarrow} \hat{B} = (X'X)^{-1} X'Y$$

A.K.A. general linear model

Model Formulation of Multivariate Linear Regression

Assume $\underset{n \times r}{Y} = \underset{n \times q}{X} \underset{q \times r}{B} + \underset{n \times r}{W}$

* If $E[W] = 0$ then $\hat{B}$ is unbiased

* If $\underset{\sim}{w}_i \overset{iid}{\sim} N(\underset{\sim}{0}, \Sigma_w)$ where $\underset{\sim}{w}_i$ are rows of W

  then:

  - $\hat{B}$ is the maximum likelihood estimator of B
  - Elements of $\hat{B}$ are normally distributed

    with $\mathbb{V}[\hat{\underset{\sim}{b}}_\ell] = \sigma_{w,\ell\ell}(X'X)^{-1}$

    $\text{Cov}[\hat{\underset{\sim}{b}}_\ell, \hat{\underset{\sim}{b}}_k] = \sigma_{w_{\ell k}}(X'X)^{-1}$

    (where $\underset{\sim}{b}_\ell$ is the ℓ-th column of B)

  etc.

* We know the distribution of $R = Y - X\hat{B}$
* Residuals R and $\hat{B}$ are independent

all of this can use to do testing

There are equivalent definitions of AIC/AICc/BIC

$$\hat{\Sigma}_w = \frac{R'R}{n}$$] maximum likelihood estimator of the noise variance

$$\text{AIC}: \ln\left(|\hat{\Sigma}_w|\right) + \frac{2}{n}\left(rk + \frac{r(r+1)}{2}\right)$$

$$\text{AICc}: \ln\left(|\hat{\Sigma}_w|\right) + \frac{r(n+q)}{n-r-q-1}$$

$$\text{BIC/SIC}: \ln\left(|\hat{\Sigma}_w|\right) + (kr + r(r+1)/2) \times \log(n)/n$$

Multivariate Time Series Concepts

Again, suppose we observe $Y = [y_1, \dots, y_r]$

We can always decompose $Y = M_Y + W$

random errors

fixed mean

Each column of Y comprised of n equally spaced observations ordered in time

Characterized by mean function $E[y_{ij}] = m_{ij}$ and

covariance function $\gamma_{ij}(s,t) = \text{Cov}[y_{si}, y_{tj}]$

when $i=j$, $\gamma_{ii}(s,t)$ is the autocovariance function of y_i

when $i \neq j$, we call $\gamma_{ij}(s,t)$ the cross-covariance function of y_i and y_j

Assume joint stationarity to simplify things in the multivariate setting

* Second moments of y_{ti} are finite for all t and i, $\mathbb{E}[y_{ti}^2] < \infty$
* The mean function is constant for each time series and doesn't depend on time

$$m_{ti} = m_i$$

* The autocovariance function $\gamma_{ii}(s,t)$ depends on s and t only through their absolute difference $\gamma_{ii}(s,t)$ can be written as $\gamma_{ii}(|s-t|)$
* The crosscovariance function $\gamma_{ij}(s,t)$ depends on s and t only through their absolute difference $\gamma_{ij}(s,t)$ can be written as $\gamma_{ij}(|s-t|)$