Bayesian Multivariate Multiple Linear Regression

Dr. Daniel B. Rowe
Professor of Computational Statistics
Department of Mathematical and Statistical Sciences
Marquette University

Outline

Univariate Simple Sampling

Univariate Simple Linear Regression

Univariate Multiple Linear Regression

Review Material Covering for Completeness.

Matrix Normal and Matrix-T PDFs

Multivariate Multiple Linear Regression

Bayesian Multivariate Linear Regression

New Material

Discussion

Homework

Reordered

Univariate Simple Sampling

Model i^{th} Observation:

$$y_i = \underbrace{\beta_0}_{\mu} + \varepsilon_i$$

$$E(\varepsilon_i) = 0 \qquad \text{cov}(\varepsilon_i) = \sigma^2$$

or equivalently

$$i = 1, \ldots, n$$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{(1)} \underset{1\times 1}{\underbrace{\left(\beta_0\right)}_{\mu}} + \underset{1\times 1}{\varepsilon_i}$$

Univariate Simple Sampling

$$y_i = \beta_0 + \varepsilon_i$$

Model n Observations:

$$i = 1, \ldots, n$$

$$\underset{n \times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n \times 1}{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}} \left(\underset{\mu}{\beta_0} \right) + \underset{n \times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

more compactly

$$\underset{n \times 1}{y} = \underset{n \times 1}{X} \; \underset{1 \times 1}{\beta} + \underset{n \times 1}{\varepsilon}$$

Univariate Simple Sampling

$y_i = \beta_0 + \varepsilon_i$

The model

column of 1's

$$\underset{n\times 1}{y} = \underset{n\times 1}{X}\ \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad E(\varepsilon)=0 \qquad \text{cov}(\varepsilon)=\sigma^2 I_n \qquad i=1,\ldots,n$$

μ

with coefficients and variance estimated as

$$\underset{1\times 1}{\hat{\beta}} = \underset{1\times 1}{(X'X)^{-1}}\ \underset{1\times n}{X'}\ \underset{n\times 1}{y} \qquad \underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta})$$

$\hat{\mu} = \bar{x}$

from likelihood function

Require $n\geq 1$, general rule $n\geq 10$

$$f(\underset{n\times 1}{y}\,|\,\beta,\sigma^2) = (2\pi)^{-n/2}\,|\sigma^2 I_{q+1}|^{-n/2}\exp\left[-\frac{1}{2}(\sigma^2 I_{q+1})^{-1}(y - X\beta)'(y - X\beta)\right]$$

.

Univariate Simple Sampling

$$\underset{n\times 1}{y} = \underset{n\times 1}{X}\ \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Using the distributional assumption that if ε is normally distributed,

$$\underset{1\times 1}{\hat{\beta}} = (X'X)^{-1}X'\ y \quad \text{is } N(\underset{n\times 1}{E(\hat{\beta})} = \beta,\ \text{cov}(\hat{\beta}) = \sigma^2 \underset{1\times 1}{(X'X)}^{-1})$$

$n\geq 1$, general rule $n\geq 10$

$\hat{\mu} = \bar{x}$

and

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-1}(y - \underset{1\times n}{X}\hat{\beta})'(y - \underset{n\times 1}{X}\hat{\beta}) \qquad \text{is gamma}(\nu,\sigma^2/\nu)!$$

$$f(\underset{n\times 1}{y}|\beta,\sigma^2) = (2\pi)^{-n/2}|\sigma^2 I_{q+1}|^{-n/2}\exp\left[-\frac{1}{2}(\sigma^2 I_{q+1})^{-1}(y - X\beta)'(y - X\beta)\right]$$

Univariate Simple Linear Regression

Model i^{th} Observation:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$E(\varepsilon_i) = 0 \qquad \operatorname{cov}(\varepsilon_i) = \sigma^2$$

$$i = 1, \ldots, n$$

or equivalently

$$\underset{1\times 1}{y_i} = \underset{1\times 2}{(1, x_i)} \underset{2\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}} + \underset{1\times 1}{\varepsilon_i}$$

Univariate Simple Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Model n Observations:

$$i = 1, \ldots, n$$

$$\underset{n\times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n\times 2}{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}} \underset{2\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}} + \underset{n\times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

more compactly

$$\underset{n\times 1}{y} = \underset{n\times 2}{X} \; \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Univariate Simple Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

The model

$i = 1,\ldots,n$

$$\underset{n\times 1}{y} = \underset{n\times 2}{X}\ \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad E(\varepsilon) = 0 \qquad \text{cov}(\varepsilon) = \sigma^2 I_n \qquad i = 1,\ldots,n$$

with coefficients and variance estimated as

$$\underset{2\times 1}{\hat{\beta}} = \underset{2\times 2}{(X'X)^{-1}}\ \underset{2\times n}{X'}\ \underset{n\times 1}{y} \qquad \underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta})$$

from likelihood function

Require $n \geq 2$, general rule $n \geq 20$

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi)^{-n/2} \left| \sigma^2 I_2 \right|^{-n/2} \exp\left[-\frac{1}{2} (\sigma^2 I_2)^{-1} (y - X\beta)'(y - X\beta) \right]$$

.

Univariate Simple Linear Regression

$$\underset{n\times 1}{y} = \underset{n\times 2}{X}\ \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Using the distributional assumption that if ε is normally distributed,

$\underset{2\times 1}{\hat{\beta}} = (X'X)^{-1}X'\ y$ is $N(\underset{2\times 1}{E(\hat{\beta})} = \beta,\ \mathrm{cov}(\hat{\beta}) = \sigma^2 (\underset{2\times 2}{X'X})^{-1})$ $n\geq 2$, general rule $n>20$

and

$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-2}(y - \underset{1\times n}{X\hat{\beta}})'(y - \underset{n\times 1}{X\hat{\beta}})$ is gamma(ν,σ^2/ν)!

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi)^{-n/2} \left|\sigma^2 I_2\right|^{-n/2} \exp\left[-\frac{1}{2}(\sigma^2 I_2)^{-1}(y - X\beta)'(y - X\beta)\right]$$

Univariate Multiple Linear Regression

Model i^{th} Observation:

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$

$$E(\varepsilon_i) = 0 \qquad \text{cov}(\varepsilon_i) = \sigma^2$$

or equivalently

$$i = 1, \ldots, n$$

$$\underset{1 \times 1}{y_i} = \underset{1 \times (q+1)}{(1, x_{i1}, \ldots, x_{iq})} \underset{(q+1) \times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}} + \underset{1 \times 1}{\varepsilon_i}$$

Univariate Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$

Model n Observations:

$$i = 1, \ldots, n$$

$$\underset{n \times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n \times (q+1)}{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}} \underset{(q+1) \times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}} + \underset{n \times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

more compactly

$$\underset{n \times 1}{y} = \underset{n \times (q+1)}{X} \underset{(q+1) \times 1}{\beta} + \underset{n \times 1}{\varepsilon}$$

Univariate Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$

The model

$$i = 1, \ldots, n$$

$$\underset{n\times 1}{y} = \underset{n\times (q+1)}{X} \underset{(q+1)\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad E(\varepsilon) = 0 \qquad \mathrm{cov}(\varepsilon) = \sigma^2 I_n \qquad i = 1, \ldots, n$$

with coefficients and variance estimated as

$$\underset{(q+1)\times 1}{\hat{\beta}} = \underset{(q+1)\times (q+1)}{(X'X)^{-1}} \underset{(q+1)\times n}{X'} \underset{n\times 1}{y} \qquad \underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta})$$

from likelihood function

Require $n \geq (q+1)$, general rule $n \geq 10(q+1)$

$$f(\underset{n\times 1}{y} | \beta, \sigma^2) = (2\pi)^{-n/2} |\sigma^2 I_{q+1}|^{-n/2} \exp\left[-\frac{1}{2}(\sigma^2 I_{q+1})^{-1}(y - X\beta)'(y - X\beta) \right]$$

.

Univariate Multiple Linear Regression

$$\underset{n\times 1}{y} = \underset{n\times 2}{X} \; \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Using the distributional assumption that if ε is normally distributed,

$\underset{(q+1)\times 1}{\hat{\beta}} = (X'X)^{-1}X'y$ is $N(\underset{(q+1)\times 1}{E(\hat{\beta})} = \beta, \ \mathrm{cov}(\hat{\beta}) = \sigma^2 \underset{(q+1)\times(q+1)}{(X'X)^{-1}})$ $\quad$ $n \geq q+1$, general rule $n > 10(q+1)$

and

$\underset{1\times 1}{\hat{\sigma}^2} = \dfrac{1}{n-q-1}(\underset{1\times n}{y - X\hat{\beta}})'(\underset{n\times 1}{y - X\hat{\beta}})$ is gamma(ν,σ^2/ν)!

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi)^{-n/2} \left| \sigma^2 I_{q+1} \right|^{-n/2} \exp\left[-\frac{1}{2} (\sigma^2 I_{q+1})^{-1} (y - X\beta)'(y - X\beta) \right]$$

Matrix Normal and Matrix-T PDFs

Separable used for vectors over time or row-col correlation in images..

Let $\underset{np\times 1}{y}$ be have a multivariate normal PDF, $\underset{np\times 1}{y} \sim N(\underset{np\times 1}{\mu}, \underset{np\times np}{\Omega})$

$$f(\underset{np\times 1}{y} | \underset{np\times 1}{\mu}, \underset{np\times np}{\Omega}) = (2\pi)^{-np/2} |\Omega|^{-1/2} \exp\left[-\frac{1}{2}(y-\mu)'\Omega^{-1}(y-\mu)\right] \qquad y,\mu \in \mathbb{R}^{np} \qquad \Omega > 0$$

and that Ω has a separable or Kronecker product structure $\underset{np\times np}{\Omega}=\underset{n\times n}{\Phi}\otimes\underset{p\times p}{\Sigma}$.
We utilize mathematical results that

$$|\underset{np\times np}{\Omega}| = |\underset{n\times n}{\Phi} \otimes \underset{p\times p}{\Sigma}| = |\underset{n\times n}{\Phi}|^{p} |\underset{p\times p}{\Sigma}|^{n} \quad \text{and} \quad (y-\mu)'(\Phi\otimes\Sigma)^{-1}(y-\mu) = tr\Sigma^{-1}(Y-M)'\Phi^{-1}(Y-M)$$

where $\underset{np\times 1}{y}=vec(\underset{n\times p}{Y})$ and $\underset{np\times 1}{\mu}=vec(\underset{n\times p}{M})$ to arrive at a reformulated PDF.

$$Y = (y_1,\ldots,y_n)$$
$$M = (\mu_1,\ldots,\mu_n)$$

Matrix Normal and Matrix-T PDFs

Separable used for vectors over time or row-col correlation in images..

The new Matrix Normal PDF for $\underset{n\times p}{Y}$ is

$$f(\underset{n\times p}{Y}|\underset{n\times p}{M},\ \underset{n\times n}{\Phi},\ \underset{p\times p}{\Sigma})=(2\pi)^{-np/2}|\Phi|^{-p/2}|\Sigma|^{-n/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}(Y-M)'\Phi^{-1}(Y-M)\right]$$

and we write $\underset{n\times p}{Y}\sim N(\underset{n\times p}{M},\underset{n\times n}{\Phi},\underset{p\times p}{\Sigma})$ aka $\underset{np\times 1}{y}\sim N(\underset{np\times 1}{\mu},\underset{np\times np}{\Phi\otimes\Sigma})$.

n=1→multivariate
n,p=1→univariate

There are various mean and covariance notations,

$E(\underset{n\times p}{Y})=\underset{n\times p}{M}$ or $E(\underset{np\times 1}{\text{vec}(Y)})=\underset{np\times 1}{\mu}$ and $cov(\underset{n\times p}{Y})=\underset{np\times np}{\Phi\otimes\Sigma}$ or $cov(\underset{np\times 1}{y})=\underset{np\times np}{\Phi\otimes\Sigma}$.

$$Y=(y_1,\ldots,y_n)$$
$$M=(\mu_1,\ldots,\mu_n)$$

Matrix Normal and Matrix-T PDFs

There is also a CLT result that MT becomes MN.

the Matrix-T distribution can be arrived at via a transformation of variable from a matrix normal to a standard matrix normal and a Wishart PDF.

$$f(\underset{n \times p}{Y} | \nu, M, \Sigma, \Phi) = k_T \frac{|\Phi|^{-\frac{p}{2}} |\Sigma|^{-\frac{n}{2}}}{\left| I_n + \frac{1}{\nu}\Phi^{-1}(Y-M)'\Sigma^{-1}(Y-M) \right|^{\frac{\nu+p}{2}}}$$

$$k_T = \frac{\prod_{j=1}^{n} \Gamma\left(\frac{\nu+p+1-j}{2}\right)}{(\nu\pi)^{\frac{np}{2}} \prod_{j=1}^{n} \Gamma\left(\frac{\nu+1-j}{2}\right)}$$

and we write $\underset{n \times p}{Y} \sim T(\nu, \underset{n \times p}{M}, \underset{n \times n}{\Phi}, \underset{p \times p}{\Sigma})$ aka $\underset{np \times 1}{y} \sim T(\underset{np \times 1}{\mu}, \underset{np \times np}{\Phi \otimes \Sigma})$.

n=1→multivariate
n,p=1→univariate

There are various mean and covariance notations,

$E(\underset{n \times p}{Y}) = \underset{n \times p}{M}$ or $E(\underset{np \times 1}{\text{vec}(Y)}) = \underset{np \times 1}{\mu}$ and $cov(\underset{n \times p}{Y}) = \frac{\nu}{\nu-2} \underset{np \times np}{\Phi \otimes \Sigma}$ or $cov(\underset{np \times 1}{y}) = \frac{\nu}{\nu-2} \underset{np \times np}{\Phi \otimes \Sigma}$.

$$Y = (y_1, \ldots, y_n)$$
$$M = (\mu_1, \ldots, \mu_n)$$

Multivariate Multiple Linear Regression

Model i^{th} Observation:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{i1} + \ldots + \beta_{qj} x_{iq} + \varepsilon_{ij}$$

$$E(\varepsilon_i) = 0 \qquad \operatorname{cov}(\varepsilon_i) = \Sigma$$

or equivalently

$$\begin{aligned} i &= 1, \ldots, n \\ j &= 1, \ldots, p \end{aligned}$$

$$\underset{1\times p}{(y_{i1}, \ldots, y_{ip})} = \underset{1\times(q+1)}{(1, x_{i1}, \ldots, x_{iq})} \underset{(q+1)\times p}{\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix}} + \underset{1\times p}{(\varepsilon_{i1}, \ldots, \varepsilon_{ip})}$$

$$\varepsilon_i = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{ip} \end{pmatrix}$$

Multivariate Multiple Linear Regression

$$(y_{i1},..,y_{ip}) = (1, x_{i1},..., x_{iq})\begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \beta_{21} & \beta_{22} & & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix} + (\varepsilon_{i1} + ... + \varepsilon_{ip})$$

$$i = 1,...,n$$

Model n Observations:

$$\underset{n\times p}{\begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{pmatrix}} = \underset{n\times(q+1)}{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}} \underset{(q+1)\times p}{\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix}} + \underset{n\times p}{\begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} & & \varepsilon_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{np} \end{pmatrix}}$$

$$i = 1,...,n$$

more compactly

$$\underset{n\times p}{Y} = \underset{n\times(q+1)}{X}\ \underset{(q+1)\times p}{B} + \underset{n\times p}{E}$$

$$\underset{(q+1)\times p}{B} = (\underbrace{\beta_1,...,\beta_p}_{\text{cols}}) = (\underbrace{b_1,...,b_n}_{\text{rows}})'$$

Multivariate Multiple Linear Regression

$$(y_{i1},..,y_{ip}) = (1,x_{i1},...,x_{iq})\begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \beta_{21} & \beta_{22} & & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix} + (\varepsilon_{i1}+...+\varepsilon_{ip})$$

$i=1,...,n$

The model

$$\underset{n\times p}{Y} = \underset{n\times (q+1)}{X} \underset{(q+1)\times p}{B} + \underset{n\times 1}{E} \qquad E(vec(E))=0 \qquad \text{cov}(vec(E)) = I_n \otimes \Sigma \qquad i=1,...,n$$

Each row of E has $\text{cov}(\varepsilon_i)=\Sigma$.

with coefficients covariance estimated as

$$\underset{(q+1)\times p}{\hat{B}} = \underset{(q+1)\times(q+1)}{(X'X)^{-1}} \underset{(q+1)\times n}{X'} \underset{n\times p}{Y} \qquad \underset{p\times p}{\hat{\Sigma}} = \frac{1}{n}(Y-X\hat{B})'(Y-X\hat{B})$$

Require $n\geq(q+1)p$, general rule $n\geq 10p(q+1)$

from likelihood function

$$f(\underset{n\times p}{Y}|B,\Sigma) = (2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}(Y-XB)'(Y-XB)\right]$$

.

$tr(\cdot)$ is trace, sum of diagonal elements

Multivariate Multiple Linear Regression

$$\underset{n\times p}{Y} = \underset{n\times(q+1)}{X}\ \underset{(q+1)\times 1}{B} + \underset{n\times p}{\varepsilon}$$

Using the distributional assumption that if ε is normally distributed,

$\underset{(q+1)\times p}{\hat{B}} = (X'X)^{-1}X'Y$ is $N\Big(E(\hat{B}) = B,\ \mathrm{cov}(vec(\hat{B})) = \big(\underset{(q+1)\times(q+1)}{X'X}\big)^{-1} \otimes \underset{p\times p}{\Sigma}\Big)$

$n \geq (q+1)$, general rule $n > 10p(q+1)$

and

$$\mathrm{cov}(\underset{\text{cols}}{\hat{\beta}_j}) = \sigma_j^2 (X'X)^{-1} \qquad \mathrm{cov}(\underset{\text{rows}}{\hat{b}_i}) = w_{ii}\Sigma \qquad W = (X'X)^{-1}$$

$\underset{p\times p}{\hat{\Sigma}} = \dfrac{1}{n-q-1}(\underset{p\times n}{Y - X\hat{B}})'(\underset{n\times p}{Y - X\hat{B}})$ is Wishart(ν,Σ/ν)!

$$\underset{p(q+1)\times p(q+1)}{\mathrm{cov}(vec(\hat{B}))} = \begin{bmatrix} w_{11}\Sigma & w_{12}\Sigma & & \\ w_{21}\Sigma & w_{22}\Sigma & & \\ & & \ddots & \\ & & & w_{(q+1),(q+1)}\Sigma \end{bmatrix}$$

$$f(\underset{n\times p}{Y} \mid B,\Sigma) = (2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}(Y - XB)'(Y - XB)\right]$$

Multivariate Multiple Linear Regression

Example:

It is believed that the heights of sons (y_s) and daughters (y_d) are linearly dependent upon the heights of their fathers (x_f) and mothers (x_m).

$$(y_{is}, y_{id}) = (1, x_{if}, x_{im}) \begin{pmatrix} \beta_{0s} & \beta_{0d} \\ \beta_{1s} & \beta_{1d} \\ \beta_{2s} & \beta_{2d} \end{pmatrix} + (\varepsilon_{is}, \varepsilon_{id}) \qquad (\varepsilon_{is}, \varepsilon_{id})' \sim N(0, \Sigma) \qquad i = 1, \ldots, n$$

Assume that the true parameter values are:

$$\underset{(q+1)\times p}{B} = \begin{pmatrix} 14.1 & 10.8 \\ 0.41 & 0.39 \\ 0.43 & 0.43 \end{pmatrix} \qquad \underset{p \times p}{\Sigma} = \begin{pmatrix} 2.250 & 1.125 \\ 1.125 & 2.250 \end{pmatrix}$$

could also imagine grandparents heights in model as x's

http://chance.amstat.org/2013/09/1-pagano/

Multivariate Multiple Linear Regression

Example:

Selected

$x_f = \{66,67,68,69,70,71,72,73,74,75\}$,

$x_m = \{60,61,62,63,64,65,66,67,68,69\}$

combinations (n=100), multiplied each with B

then added correlated normal noise with Σ.

Repeated L=10,000 times.

$$B = \begin{pmatrix} 14.1 & 10.8 \\ 0.41 & 0.39 \\ 0.43 & 0.43 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 2.250 & 1.125 \\ 1.125 & 2.250 \end{pmatrix}$$

$$A = \begin{pmatrix} 1.50 & 0 \\ 0.755 & 1.30 \end{pmatrix}$$

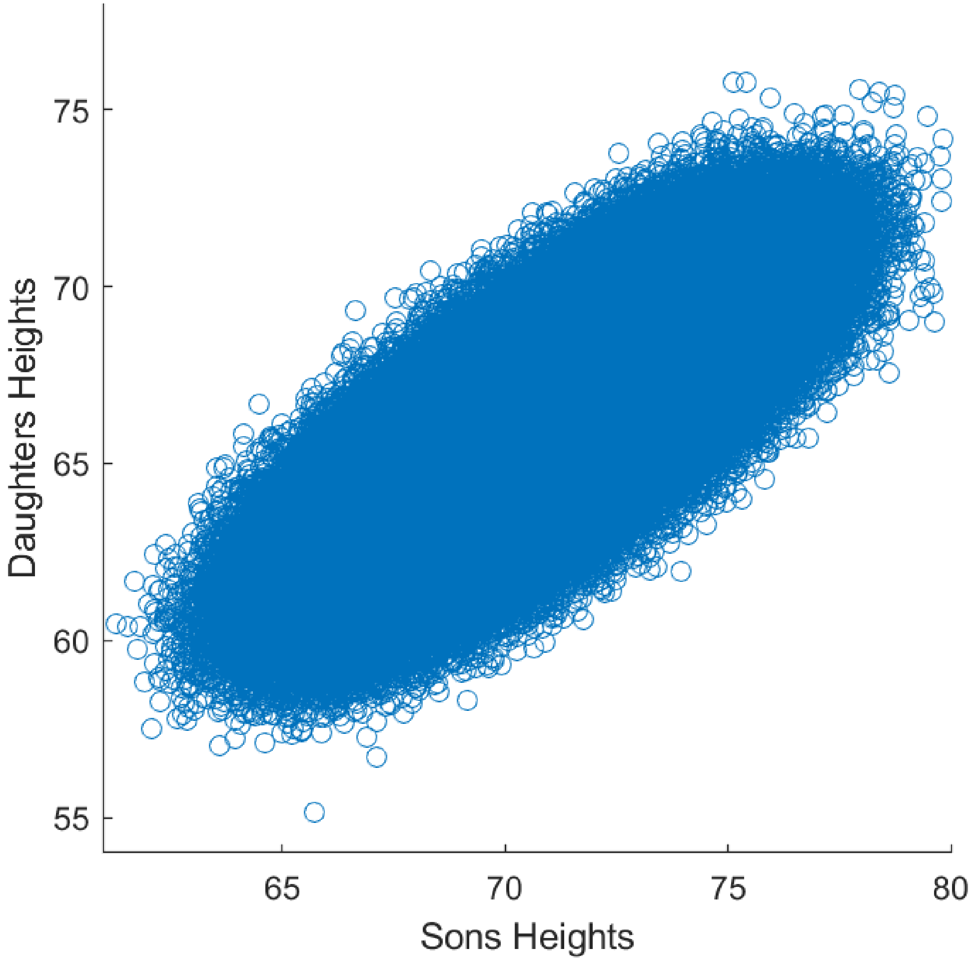
Multivariate Multiple Linear Regression

Example:

Estimates and plots from all 10^6.

$$\hat{B} = \begin{pmatrix} 14.0511 & 10.8487 \\ 0.4108 & 0.3899 \\ 0.4299 & 0.4294 \end{pmatrix}$$

$$\hat{\Sigma} = \begin{pmatrix} 2.2504 & 1.1207 \\ 1.1207 & 2.2465 \end{pmatrix}$$



$$B = \begin{pmatrix} 14.1 & 10.8 \\ 0.41 & 0.39 \\ 0.43 & 0.43 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 2.250 & 1.125 \\ 1.125 & 2.250 \end{pmatrix}$$

$$A = \begin{pmatrix} 1.50 & 0 \\ 0.755 & 1.30 \end{pmatrix}$$

Multivariate Multiple Linear Regression

```
% Multivariate model for
% Y1=son height, Y2=daughter height
% X1 = father height, X2=mother height
clear all
close all
rng('default')

% define true parameter values
n=100;
B=[[14.1;.41;.43],[10.8;.39;.43]]
sigma1=1.5; sigma2=1.5; rho=.5;
Sigma=[sigma1^2,sigma1*sigma2*rho;...
       sigma1*sigma2*rho,sigma2^2]
A=chol(Sigma)'

% generate simulated data
hf=(66:66+9);,hm=(60:60+9);
[Xf,Xm]=meshgrid(hf,hm);
xf=reshape(Xf,n,1);,xm=reshape(Xm,n,1);
X=[ones(n,1),xf,xm]
```

```
replicate=10000;
X=repmat(X,replicate,1);
n=size(X,1);

E=(A*randn(2,n))';
Y=X*B+E;

% Estimate parameters
Bhat=inv(X'*X)*X'*Y;
SigmaHat=(Y-X*B)'*(Y-X*B)/(n-3);

[Bhat,B]
[SigmaHat,Sigma]

figure;
scatter(Y(:,1),Y(:,2))
axis square
xlabel('Sons Heights')
ylabel('Daughters Heights')
xlim([61,80]),ylim([54,78])
```

Bayesian Multivariate Multiple Linear Regression

The model

$$\underset{n\times p}{Y} = \underset{n\times (q+1)}{X}\ \underset{(q+1)\times p}{B} + \underset{n\times 1}{E} \qquad E(vec(E))=0 \qquad \mathrm{cov}(vec(E)) = I_n \otimes \Sigma \qquad i=1,\ldots,n$$

Each row of E has $\mathrm{cov}(\varepsilon_i)=\Sigma$.

with likelihood

$$\underset{n\times p}{f}(Y|B,\Sigma) = (2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}(Y-XB)'(Y-XB)\right]$$

matrix normal

and priors

$$\underset{(q+1)\times p}{f}(B|\Sigma) = (2\pi)^{-p(q+1)/2}|\Sigma/n_0|^{-(q+1)/2}\exp\left[-\frac{n_0}{2}tr\Sigma^{-1}(B-B_0)'(B-B_0)\right]$$

matrix normal

$$\underset{p\times p}{f}(\Sigma) = k|H|^{\nu/2}|\Sigma|^{-(\nu+p+1)/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}H\right]$$

Inverse Wishart

Bayesian Multivariate Multiple Linear Regression

$$H_* = (Y - XB)'(Y - XB) + n_0(B - B_0)'(B - B_0) + H$$
$$\nu_* = \nu + n + q + 1$$

yields posterior

$$\underset{(q+1)\times p}{f(B,\Sigma|Y)} \propto |\Sigma|^{-(\nu_*+p+1)/2} \exp\left\{-\frac{1}{2} tr\Sigma^{-1} H_*\right\}$$

$$\hat{B} = (X'X + n_0 I)^{-1}(X'Y + n_0 B_0)$$

complete the square

$$\underset{(q+1)\times p}{f(B,\Sigma|Y)} \propto |\Sigma|^{-(\nu_*+p+1)/2} \exp\left\{-\frac{1}{2} tr\Sigma^{-1}\left[(B-\hat{B})'(X'X + n_0 I)(B-\hat{B}) + W\right]\right\}.$$

$$W = n_0 B_0' B_0 + H + Y'Y - \hat{B}'(X'X + n_0 I)\hat{B}$$

Integrate wrt to Σ to get Matrix Student-T PDF for B.

Integrate wrt to B to get Inverse Wishart PDF for Σ.

Bayesian Multivariate Multiple Linear Regression

$$f(Y|B,\Sigma)=(2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}(Y-XB)'(Y-XB)\right]$$

$$f(B|\Sigma)=(2\pi)^{-p(q+1)/2}|\Sigma/n_0|^{-(q+1)/2}\exp\left[-\frac{n_0}{2}tr\Sigma^{-1}(B-B_0)'(B-B_0)\right]$$

$$f(\Sigma)=k|H|^{\nu/2}|\Sigma|^{-(\nu+p+1)/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}H\right]$$

The marginal PDF of B can be found

$$\underset{(q+1)\times p}{f(B|Y)}=\int_\Sigma f(B,\Sigma|Y)d\Sigma$$

$$\underset{(q+1)\times p}{f(B|Y)}\propto\int_\Sigma|\Sigma|^{-(\nu_*+p+1)/2}\exp\left\{-\frac{1}{2}tr\Sigma^{-1}\left[(B-\hat{B})'(X'X+n_0I)(B-\hat{B})+W\right]\right\}d\Sigma$$

$$\nu_*=\nu+n+q+1$$

$$\underset{(q+1)\times p}{f(B|Y)}\propto|H_*|^{-\nu_*/2}\underbrace{\int_\Sigma|H_*|^{\nu_*/2}|\Sigma|^{-(\nu_*+p+1)/2}e^{-\frac{1}{2}tr\Sigma^{-1}H_*}d\Sigma}_{=\text{constant}}$$

$$H_*=(B-\hat{B})'(X'X+n_0I)(B-\hat{B})+W$$

$$W=n_0B_0'B_0+H+Y'Y-\hat{B}'(X'X+n_0I)\hat{B}$$

Bayesian Multivariate Multiple Linear Regression

$$f(Y|B,\Sigma) = (2\pi)^{-np/2} |\Sigma|^{-n/2} \exp\left[-\frac{1}{2} tr\Sigma^{-1}(Y-XB)'(Y-XB)\right]$$

$$f(B|\Sigma) = (2\pi)^{-p(q+1)/2} |\Sigma / n_0|^{-(q+1)/2} \exp\left[-\frac{n_0}{2} tr\Sigma^{-1}(B-B_0)'(B-B_0)\right]$$

$$f(\Sigma) = k |H|^{\nu/2} |\Sigma|^{-(\nu+p+1)/2} \exp\left[-\frac{1}{2} tr\Sigma^{-1} H\right]$$

The marginal PDF of μ can be found

$$\underset{(q+1)\times p}{f(B|Y)} \propto |H_*|^{-\nu_*/2}$$

$$\underset{(q+1)\times p}{f(B|Y)} \propto \frac{1}{|W+(B-\hat{B})'(X'X+n_0 I)(B-\hat{B})|^{\nu_*/2}}$$

Note: If A $p\times q$ and B is $q\times p$, then

$$|I_p + AB| = |I_q + BA|$$

$$\nu_* = \nu + n + q + 1$$
$$\nu_\# = \nu + n - p + q + 1$$

$$\underset{(q+1)\times p}{f(B|Y)} \propto \frac{1}{\left|I_{q+1} + \frac{1}{\nu_\#}(X'X+n_0 I)(B-\hat{B})T^{-1}(B-\hat{B})'\right|^{(\nu_\#+p)/2}}$$

Matrix Student-T

$$\underset{(q+1)\times p}{E(B|Y)} = \hat{B}$$

$$T = \frac{W}{\nu_\#}$$

$$H_* = (B-\hat{B})'(X'X+n_0 I)(B-\hat{B}) + W$$

$$W = n_0 B_0' B_0 + H + Y'Y - \hat{B}'(X'X+n_0 I)\hat{B}$$

Bayesian Multivariate Multiple Linear Regression

$$f(Y|B,\Sigma)=(2\pi)^{-np/2}|\Sigma|^{-n/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}(Y-XB)'(Y-XB)\right]$$

$$f(B|\Sigma)=(2\pi)^{-p(q+1)/2}|\Sigma/n_0|^{-(q+1)/2}\exp\left[-\frac{n_0}{2}tr\Sigma^{-1}(B-B_0)'(B-B_0)\right]$$

$$f(\Sigma)=k|H|^{\nu/2}|\Sigma|^{-(\nu+p+1)/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}H\right]$$

The marginal PDF of Σ can be found

$$\underset{p\times p}{f(\Sigma|Y)}=\int_B f(B,\Sigma|Y)dB$$

$$\underset{p\times p}{f(\Sigma|Y)}\propto\int_B|\Sigma|^{-(\nu_*+p+1)/2}\exp\left\{-\frac{1}{2}tr\Sigma^{-1}\left[(B-\hat{B})'(X'X+n_0I)(B-\hat{B})+W\right]\right\}dB$$

$$\nu_*=\nu+n+q+1$$

$$\underset{p\times p}{f(\Sigma|Y)}\propto|\Sigma|^{-(\nu_*+p-q)/2}e^{-\frac{1}{2}tr\Sigma^{-1}W}\underbrace{\int_B|\Sigma|^{-(q+1)/2}\exp\left\{-\frac{1}{2}tr\Sigma^{-1}(B-\hat{B})'(X'X+n_0I)(B-\hat{B})\right\}dB}_{=\text{constant}}$$

$$W=n_0B_0'B_0+H+Y'Y-\hat{B}'(X'X+n_0I)\hat{B}$$

Bayesian Multivariate Multiple Linear Regression

The marginal PDF of Σ can be found to be inverse Wishart

$$f(\underset{p\times p}{\Sigma} | Y) = k |W|^{(\nu_{**})/2} |\Sigma|^{-(\nu_{**}+p+1)/2} e^{-\frac{1}{2} tr\Sigma^{-1}W}$$

with expectation, variances, and covariances

$$E(\underset{p\times p}{\Sigma} | Y) = \frac{W}{\nu_{**}-p-1} \quad ,$$

$$\mathrm{var}(\Sigma_{ii} | Y) = \frac{2W_{ii}^2}{(\nu_{**}-p-1)^2(\nu_{**}-p-3)}$$

$$\nu_{**} = \nu + n$$

$$\mathrm{var}(\Sigma_{ij} | Y) = \frac{(\nu_{**}-p+1)W_{ij}^2 + (\nu_{**}-p-1)W_{ii}W_{jj}}{(\nu_{**}-p)(\nu_{**}-p-1)^2(\nu_{**}-p-3)} \quad , \quad \mathrm{cov}(\Sigma_{ij}, \Sigma_{kl} | Y) = \frac{2W_{ij}W_{kl} + (\nu_{**}-p-1)(W_{ik}W_{jl} + W_{il}W_{kj})}{(\nu_{**}-p)(\nu_{**}-p-1)^2(\nu_{**}-p-3)} \; .$$

$$W = n_0 B_0' B_0 + H + Y'Y - \hat{B}'(X'X + n_0 I)\hat{B}$$

Discussion

We started with the univariate simple normal samples model and built up to the multivariate multiple linear regression model.

We formulated a Bayesian multivariate multiple regression model with assessed hyperparameters for conjugate priors and estimated the model parameters exactly *a posteriori*.

Discussion

Questions?

$$\underset{n\times p}{Y} = \underset{n\times(q+1)}{X}\ \underset{(q+1)\times p}{B} + \underset{n\times p}{E}$$

Non-Bayesian

$$\underset{(q+1)\times p}{\hat{B}_{MLE}} = (X'X)^{-1}X'\ Y$$

$$\underset{(q+1)p\ \times\ (q+1)p}{\mathrm{cov}(vec(\hat{B}))} = \left(X'X\right)^{-1} \otimes \hat{\Sigma}$$

$$\hat{\Sigma} = \frac{1}{n-q-1}(Y'Y - \hat{B}'X'X\hat{B})$$

Bayesian

$$\underset{(q+1)\times p}{\hat{B}} = (X'X + n_0 I)^{-1}[X'X\hat{B}_{MLE} + n_0 B_0]$$

$$\underset{(q+1)p\ \times\ (q+1)p}{\mathrm{cov}(vec(\hat{B}))} = \left(X'X + n_0 I\right)^{-1} \otimes \frac{\nu_\#}{\nu_\# - 2} T$$

$$T = [n_0 B_0' B_0 + H + Y'Y - \hat{B}'(X'X + n_0 I)\hat{B}] / \nu_\#$$

$$\nu_\# = \nu + n - p + q + 1$$

Homework 12

1***.Select observed known father x_{if} and mother x_{im} heights, multiply by coefficients B, add correlated error correlated ε_i with mean 0 and covariance Σ, repeat n=100 times.

$$\underset{(q+1)\times p}{\hat{B}} = (\underset{(q+1)\times (q+1)}{X'X})^{-1} \underset{(q+1)\times n}{X'} \underset{n\times p}{Y}$$

$$(y_{is}, y_{id}) = (1, x_{if}, x_{im}) \begin{pmatrix} \beta_{0s} & \beta_{0d} \\ \beta_{1s} & \beta_{1d} \\ \beta_{2s} & \beta_{2d} \end{pmatrix} + (\varepsilon_{is}, \varepsilon_{id}) \qquad i = 1,\ldots,n$$

$$\underset{p\times p}{\hat{\Sigma}} = \frac{(Y - X\hat{B})'(Y - X\hat{B})}{n - q - 1}$$

form Y and X, estimate $\hat{B}$ and $\hat{\Sigma}$.

Repeat a large number of times L so that you have $\hat{B}^{(1)},\ldots,\hat{B}^{(L)}$ and $\hat{\Sigma}^{(1)},\ldots,\hat{\Sigma}^{(L)}$.

Compute means, variances, make histograms, etc.

Compare to true values.

*** For enthusiastic students.

Homework 12

2***.Assess hyperparameters B_0, n_0, v, and H for the conjugate prior PDFs.

Compute Bayesian posterior estimates of data sets generated in #2.

try

$$B_0 = \begin{pmatrix} 14 & 11 \\ 0.4 & 0.4 \\ 0.4 & 0.4 \end{pmatrix}$$

*** For enthusiastic students.

Univariate Simple Sampling

Slides Reordered

Model i^{th} Observation:

$$y_i = \overset{\mu}{\beta_0} + \varepsilon_i \qquad E(\varepsilon_i) = 0 \qquad \operatorname{cov}(\varepsilon_i) = \sigma^2$$

or equivalently $\qquad i = 1, \ldots, n$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{(1)} \underset{1\times 1}{\left(\overset{\mu}{\beta_0}\right)} + \underset{1\times 1}{\varepsilon_i}$$

Univariate Simple Linear Regression

Model i^{th} Observation:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad E(\varepsilon_i) = 0 \qquad \operatorname{cov}(\varepsilon_i) = \sigma^2$$

or equivalently $\qquad i = 1, \ldots, n$

$$\underset{1\times 1}{y_i} = \underset{1\times 2}{(1, x_i)} \underset{2\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}} + \underset{1\times 1}{\varepsilon_i}$$

Univariate Multiple Linear Regression

Model i^{th} Observation:

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$

$$E(\varepsilon_i) = 0 \qquad \operatorname{cov}(\varepsilon_i) = \sigma^2$$

$$i = 1, \ldots, n$$

or equivalently

$$\underset{1\times 1}{y_i} = \underset{1\times (q+1)}{(1, x_{i1}, \ldots, x_{iq})} \underset{(q+1)\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}} + \underset{1\times 1}{\varepsilon_i}$$

Multivariate Multiple Linear Regression

Model i^{th} Observation:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{i1} + \ldots + \beta_{qj} x_{iq} + \varepsilon_{ij}$$

$$E(\varepsilon_i) = 0 \qquad \text{cov}(\varepsilon_i) = \Sigma$$

or equivalently

$$i = 1, \ldots, n$$
$$j = 1, \ldots, p$$

$$\underset{1 \times p}{(y_{i1}, .., y_{ip})} = \underset{1 \times (q+1)}{(1, x_{i1}, \ldots, x_{iq})} \underset{(q+1) \times p}{\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix}} + \underset{1 \times p}{(\varepsilon_{i1}, \ldots, \varepsilon_{ip})}$$

$$\varepsilon_i = \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{ip} \end{pmatrix}$$

Univariate Simple Sampling

$$y_i = \beta_0 + \varepsilon_i$$

Model n Observations:

$$i = 1, \ldots, n$$

$$\underset{n\times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n\times 1}{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}} \left(\underset{\mu}{\beta_0} \right) + \underset{n\times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

more compactly

$$\underset{n\times 1}{y} = \underset{n\times 1}{X} \; \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Univariate Simple Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Model n Observations:

$$i = 1, \ldots, n$$

$$\underset{n\times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n\times 2}{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}} \underset{2\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}} + \underset{n\times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

more compactly

$$\underset{n\times 1}{y} = \underset{n\times 2}{X} \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Univariate Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$

Model n Observations:

$$i = 1, \ldots, n$$

$$\underset{n \times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n \times (q+1)}{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}} \underset{(q+1) \times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}} + \underset{n \times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

more compactly

$$\underset{n \times 1}{y} = \underset{n \times (q+1)}{X} \underset{(q+1) \times 1}{\beta} + \underset{n \times 1}{\varepsilon}$$

Multivariate Multiple Linear Regression

$$(y_{i1},..,y_{ip}) = (1, x_{i1},...,x_{iq}) \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \beta_{21} & \beta_{22} & & \beta_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix} + (\varepsilon_{i1} + ... + \varepsilon_{ip})$$

$$i = 1,...,n$$

Model n Observations:

$$\underset{n \times p}{\begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{pmatrix}} = \underset{n \times (q+1)}{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}} \underset{(q+1) \times p}{\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix}} + \underset{n \times p}{\begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} & & \varepsilon_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{np} \end{pmatrix}}$$

$$i = 1,...,n$$

more compactly

$$\underset{n \times p}{Y} = \underset{n \times (q+1)}{X} \; \underset{(q+1) \times p}{B} + \underset{n \times p}{E}$$

$$\underset{(q+1) \times p}{B} = (\underbrace{\beta_1,...,\beta_p}_{\text{cols}}) = (\underbrace{b_1,...,b_n}_{\text{rows}})'$$

Univariate Simple Sampling

$$y_i = \beta_0 + \varepsilon_i$$

The model

$$\underset{n\times 1}{y} = \underset{n\times 1}{X} \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad E(\varepsilon)=0 \qquad \text{cov}(\varepsilon)=\sigma^2 I_n \qquad i=1,\ldots,n$$

μ

with coefficients and variance estimated as

$$\underset{1\times 1}{\hat{\beta}} = \underset{1\times 1}{(X'X)^{-1}} \underset{1\times n}{X'} \underset{n\times 1}{y} \qquad \underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta})$$

$\hat{\mu} = \bar{x}$

Require $n\geq 1$, general rule $n\geq 10$

from likelihood function

$$f(\underset{n\times 1}{y}|\beta,\sigma^2) = (2\pi)^{-n/2}|\sigma^2 I_{q+1}|^{-n/2}\exp\left[-\frac{1}{2}(\sigma^2 I_{q+1})^{-1}(y - X\beta)'(y - X\beta)\right]$$

.

Univariate Simple Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$i = 1, \ldots, n$$

The model

$$\underset{n\times 1}{y} = \underset{n\times 2}{X} \; \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad E(\varepsilon) = 0 \qquad \operatorname{cov}(\varepsilon) = \sigma^2 I_n \qquad i = 1, \ldots, n$$

with coefficients and variance estimated as

$$\underset{2\times 1}{\hat{\beta}} = \underset{2\times 2}{(X'X)^{-1}} \; \underset{2\times n}{X'} \; \underset{n\times 1}{y} \qquad \underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta})$$

from likelihood function

Require $n \geq 2$, general rule $n \geq 20$

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi)^{-n/2} \left| \sigma^2 I_2 \right|^{-n/2} \exp\left[-\frac{1}{2} (\sigma^2 I_2)^{-1} (y - X\beta)'(y - X\beta) \right]$$

.

Univariate Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$

The model

$i = 1, \ldots, n$

$$\underset{n\times 1}{y} = \underset{n\times (q+1)}{X}\ \underset{(q+1)\times 1}{\beta} + \underset{n\times 1}{\varepsilon} \qquad E(\varepsilon) = 0 \qquad \mathrm{cov}(\varepsilon) = \sigma^2 I_n \qquad i = 1, \ldots, n$$

with coefficients and variance estimated as

$$\underset{(q+1)\times 1}{\hat{\beta}} = \underset{(q+1)\times(q+1)}{(X'X)^{-1}}\ \underset{(q+1)\times n}{X'}\ \underset{n\times 1}{y} \qquad \underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n}(y - X\hat{\beta})'(y - X\hat{\beta})$$

Require $n \geq (q+1)$, general rule $n \geq 10(q+1)$

from likelihood function

$$f(\underset{n\times 1}{y} | \beta, \sigma^2) = (2\pi)^{-n/2} |\sigma^2 I_{q+1}|^{-n/2} \exp\left[-\frac{1}{2}(\sigma^2 I_{q+1})^{-1}(y - X\beta)'(y - X\beta)\right]$$

.

Multivariate Multiple Linear Regression

$$(y_{i1},..,y_{ip})=(1,x_{i1},...,x_{iq})\begin{pmatrix}\beta_{11} & \beta_{12} & \cdots & \beta_{1p}\\ \beta_{21} & \beta_{22} & & \beta_{2p}\\ \vdots & \vdots & \ddots & \vdots\\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp}\end{pmatrix}+(\varepsilon_{i1}+...+\varepsilon_{ip})$$

$$i=1,...,n$$

The model

$$\underset{n\times p}{Y}=\underset{n\times (q+1)}{X}\ \underset{(q+1)\times p}{B}+\underset{n\times 1}{E} \qquad E(vec(E))=0 \qquad \text{cov}(vec(E))=I_n\otimes\Sigma \qquad i=1,...,n$$

Each row of E has $\text{cov}(\varepsilon_i)=\Sigma$.

with coefficients covariance estimated as

$$\underset{(q+1)\times p}{\hat{B}}=\underset{(q+1)\times(q+1)}{(X'X)^{-1}}\ \underset{(q+1)\times n}{X'}\ \underset{n\times p}{Y} \qquad \underset{p\times p}{\hat{\Sigma}}=\frac{1}{n}(Y-X\hat{B})'(Y-X\hat{B})$$

Require $n\geq(q+1)p$, general rule $n\geq 10p(q+1)$

from likelihood function

$$f(\underset{n\times p}{Y}\,|\,B,\Sigma)=(2\pi)^{-np/2}\,|\Sigma|^{-n/2}\exp\left[-\frac{1}{2}tr\Sigma^{-1}(Y-XB)'(Y-XB)\right]$$

.

$tr(\cdot)$ is trace, sum of diagonal elements

Univariate Simple Sampling

$$\underset{n\times 1}{y} = \underset{n\times 1}{X}\ \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Using the distributional assumption that if ε is normally distributed,

$\underset{1\times 1}{\hat{\beta}} = (X'X)^{-1}X'\ y$ is $N(\underset{n\times 1}{E(\hat{\beta})} = \beta,\ \mathrm{cov}(\hat{\beta}) = \sigma^2 \underset{1\times 1}{(X'X)^{-1}})$ $n\geq 1$, general rule $n>10$

$\hat{\mu} = \bar{x}$

and

$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-1}(y - X\hat{\beta})'(y - X\hat{\beta})$ is gamma(ν,σ^2/ν)!

$$f(\underset{n\times 1}{y} | \beta, \sigma^2) = (2\pi)^{-n/2} |\sigma^2 I_{q+1}|^{-n/2} \exp\left[-\frac{1}{2}(\sigma^2 I_{q+1})^{-1}(y - X\beta)'(y - X\beta)\right]$$

Univariate Simple Linear Regression

$$\underset{n\times 1}{y} = \underset{n\times 2}{X}\ \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Using the distributional assumption that if ε is normally distributed,

$$\underset{2\times 1}{\hat{\beta}} = (X'X)^{-1}X'\ y \quad \text{is } N(\underset{2\times 1}{E(\hat{\beta})} = \beta,\ \mathrm{cov}(\hat{\beta}) = \sigma^2 (\underset{2\times 2}{X'X})^{-1})$$

$n \geq 2$, general rule $n > 20$

and

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-2}\underset{1\times n}{(y - X\hat{\beta})'}\underset{n\times 1}{(y - X\hat{\beta})} \qquad \text{is gamma}(\nu, \sigma^2/\nu)!$$

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi)^{-n/2} |\sigma^2 I_2|^{-n/2} \exp\left[-\frac{1}{2}(\sigma^2 I_2)^{-1}(y - X\beta)'(y - X\beta)\right]$$

Univariate Multiple Linear Regression

$$\underset{n\times 1}{y} = \underset{n\times 2}{X}\ \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Using the distributional assumption that if ε is normally distributed,

$$\underset{(q+1)\times 1}{\hat{\beta}} = (X'X)^{-1}X'\ y \quad \text{is } N(\underset{(q+1)\times 1}{E(\hat{\beta}) = \beta}, \ \text{cov}(\hat{\beta}) = \sigma^2 \underset{(q+1)\times(q+1)}{(X'X)^{-1}})$$

$n \geq 2$, general rule $n > 20$

and

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-q-1}\underset{1\times n}{(y - X\hat{\beta})'}\underset{n\times 1}{(y - X\hat{\beta})} \text{ is gamma}(\nu,\sigma^2/\nu)!$$

$$f(\underset{n\times 1}{y}|\beta,\sigma^2) = (2\pi)^{-n/2}|\sigma^2 I_{q+1}|^{-n/2}\exp\left[-\frac{1}{2}(\sigma^2 I_{q+1})^{-1}(y - X\beta)'(y - X\beta)\right]$$

Multivariate Multiple Linear Regression

$$\underset{n\times p}{Y} = \underset{n\times(q+1)}{X}\ \underset{(q+1)\times 1}{B} + \underset{n\times p}{\varepsilon}$$

Using the distributional assumption that if ε is normally distributed,

$\underset{(q+1)\times p}{\hat{B}} = (X'X)^{-1}X'\,Y$ is $N\left(E(\hat{B}) = B,\ \mathrm{cov}(vec(\hat{B})) = \left(\underset{(q+1)\times(q+1)}{X'X}\right)^{-1} \otimes \underset{p\times p}{\Sigma}\right)$ $\quad n \geq (q+1)$, general rule $n > 10p(q+1)$

and $\quad \underset{\text{cols}}{\mathrm{cov}(\hat{\beta}_j)} = \sigma_j^2 \left(X'X\right)^{-1} \quad \underset{\text{rows}}{\mathrm{cov}(\hat{b}_i)} = w_{ii}\Sigma \quad W = \left(X'X\right)^{-1}$

$\underset{p\times p}{\hat{\Sigma}} = \dfrac{1}{n-q-1}(Y - \underset{p\times n}{X\hat{B}})'(Y - \underset{n\times p}{X\hat{B}})$ is Wishart(ν,Σ/ν)!

$$\mathrm{cov}(vec(\hat{B})) = \begin{bmatrix} w_{11}\Sigma & w_{12}\Sigma & & \\ w_{21}\Sigma & w_{22}\Sigma & & \\ & & \ddots & \\ & & & w_{(q+1),(q+1)}\Sigma \end{bmatrix}$$

$$f(\underset{n\times p}{Y} \mid B,\Sigma) = (2\pi)^{-np/2}\,|\Sigma|^{-n/2} \exp\left[-\frac{1}{2} tr\Sigma^{-1}(Y - XB)'(Y - XB)\right]$$