de densidad conjunta de $Y_1, \ldots, Y_N$ condicional a $X_1 = x_1; \ldots; X_N = x_N$ es

$$L(\beta_0, \beta_1, \sigma^2) = f_{Y_1 \ldots Y_N | X_1 \ldots X_N}(y_1, \ldots, y_N | x_1, \ldots, x_N, \beta_0, \beta_1, \sigma^2)$$

$$= \prod_{i=1}^{N} f_{Y_i|X_i}(y_i | x_i, \beta_0, \beta_1, \sigma^2)$$

$$= \prod_{i=1}^{N} \left\{ (2\pi\sigma^2)^{-1/2} e^{-\frac{1}{2\sigma^2}\left[y_i - (\beta_0 + \beta_1 x_i)\right]^2} \right\}$$

$$= (2\pi\sigma^2)^{-N/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \beta_0 - \beta_1 x_i)^2}$$

$$l(\beta_0, \beta_1, \sigma^2) = \operatorname{Log} L(\beta_0, \beta_1, \sigma^2)$$

$$= -\frac{N}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \beta_0 - \beta_1 x_i)^2$$

Es claro que maximizar $l(\beta_0, \beta_1, \sigma^2)$ c.r. a β_0 y β_1 corresponde con minimizar la suma de cuadrados

$$\sum_{i=1}^{N}(y_i - \beta_0 - \beta_1 x_i)^2 \qquad \left(= \Delta(\beta_0, \beta_1)\right)$$

En consecuencia, los estimadores de máxima verosimilitud para β_0 y β_1 coinciden con los estimadores de mínimos cuadrados. El sistema de ecuaciones que se obtiene cuando $\frac{\partial l}{\partial \beta_0}$ y $\frac{\partial l}{\partial \beta_1}$ se igualan a 0, es el mismo que analizamos cuando estabamos buscando los estimadores de mínimos cuadrados.

Por esta razón el sistema de ecuaciones (EN) en la página 36 se llama sistema Normal (ó ecuaciones normales). En cuanto al tercer parámetro σ^2, la derivada parcial correspondiente es

$$\frac{\partial l}{\partial \sigma^2} = -\frac{N}{2}\frac{2\pi}{2\pi\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{N}(y_i-\beta_0-\beta_1 x_i)^2 .$$

Igualando a 0 obtenemos

$$-\frac{n}{\hat{\sigma}^2} + \frac{1}{\hat{\sigma}^4}\sum_{i=1}^{N}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_i)^2 ,$$

de donde

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_i)^2$$

En resumen, los estimadores de máxima verosimilitud $\hat{\theta}^{(mv)} = (\hat{\beta}_0^{(mv)}, \hat{\beta}_1^{(mv)}, \hat{\sigma}^{2(mv)})$ para $\theta = (\beta_0, \beta_1, \sigma^2)$ están dados por

$$\hat{\beta}_1^{(mv)} = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{N}(x_i-\bar{x})^2},$$

$$\hat{\beta}_0^{(mv)} = \bar{y} - \hat{\beta}_1\bar{x},$$

y $$\hat{\sigma}^{2(mv)} = \frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_i)^2,$$

donde $\hat{\beta}_0^{(mv)}$ y $\hat{\beta}_1^{(mv)}$ coinciden con los estimadores de mínimos cuadrados para β_0 y β_1. Debido a lo anterior podemos escribir

$$\hat{\beta}_1^{(mv)} = \sum_{i=1}^{N} a_i Y_i \qquad \hat{\beta}_0^{(mv)} = \sum_{i=1}^{N} b_i Y_i \text{ (1)} \cdots\cdots (1)$$

y además sabemos que

$$\mathbb{E}(\hat{\beta}_1^{(mv)}) = \beta_1 \quad \text{y} \quad VAR(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{N}(x_i-\bar{x})^2} \cdots\cdots (2)$$

Además

$$\mathbb{E}(\hat{\beta}_0^{(mv)}) = \beta_0 \quad \text{y} \quad VAR(\hat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^{N} x_i^2}{N\sum_{i=1}^{N}(x_i-\bar{x})^2} \cdots\cdots (3)$$

(1) donde $a_i = \frac{x_i-\bar{x}}{\sum_{i}^{N}(x_i-\bar{x})^2}$ y $b_i = \frac{1}{N} - a_i\bar{x}$

Usando las relaciones (1), (2) y (3)[(1)], podemos deducir que

$$\hat{\beta}_1^{(MV)} \sim \text{Normal}\left(\beta_1, \frac{\sigma^2}{\sum_{i=1}^{N}(x_i-\bar{x})^2}\right)$$

$$\hat{\beta}_0^{(MV)} \sim \text{Normal}\left(\beta_0, \frac{\sigma^2 \sum_{i=1}^{N} x_i^2}{N\sum_{i=1}^{N}(x_i-\bar{x})^2}\right)$$

<u>NOTA</u>: $\hat{\beta}_0^{(MV)}$ y $\hat{\beta}_1^{(MV)}$ vistos como v.a. <u>no</u> necesariamente son independientes:

$$\begin{aligned}
\text{COV}(\hat{\beta}_0, \hat{\beta}_1) &= \text{COV}\left(\sum_{i=1}^{N} b_i Y_i, \sum_{i=1}^{N} a_i Y_i\right) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{N} a_i b_j \,\text{COV}(Y_i, Y_j) \\
&\overset{(1)}{=} \sum_{i=1}^{N} a_i b_i \,\text{VAR}(Y_i) \\
&= \sigma^2 \sum_{i=1}^{N} a_i b_i \\
&= \sigma^2 \sum_{i=1}^{N} a_i\left(\frac{1}{N} - \bar{x} a_i\right) \\
&= \sigma^2\left[\frac{1}{N}\sum_{i=1}^{N} a_i - \bar{x}\sum_{i=1}^{N} a_i^2\right] = -\sigma^2 \frac{\bar{x}}{\sum_{i=1}^{N}(x_i-\bar{x})^2}
\end{aligned}$$

(1) Y tomando en cuenta que $Y_i|X_i=x_i \sim \text{Normal}(\beta_0+\beta_1 x_i, \sigma^2)$ $i=1,2,\dots,N$, con $f_{Y_1\cdots Y_N|X_1=x_1\cdots X_N=x_N} = \prod_{i=1}^{N} f_{Y_i|X_i=x_i}$ y con $X_1,\dots,X_N$ v.a. independ. ($\Rightarrow Y_1\cdots Y_N$ v.a. independ.)

★ **Proposición**: Si $\mathbb{Y}' = (Y_1, \dots, Y_N)$ es un vector aleatorio tal que $\mathbb{Y} \sim N_N(\mu_{\mathbb{Y}}; \Sigma_{\mathbb{Y}})$ y $\mathbb{A}$ es una matriz de dimensiones $p \times N$ entonces

$$\mathbb{X} \equiv \mathbb{A}\mathbb{Y} \sim N_p(\mu_X, \Sigma_X) \quad \text{donde}$$

$$\mu_X = \mathbb{A}\mu_{\mathbb{Y}} \quad y \quad \Sigma_X = \mathbb{A}\Sigma_Y \mathbb{A}'.$$

Usaremos este resultado con $\mathbb{Y}' = (Y_1, \dots, Y_N)$,

$$\mu_{\mathbb{Y}}' = (\beta_0 + \beta_1 x_1, \dots, \beta_0 + \beta_1 x_N) \; ; \; \Sigma_{\mathbb{Y}} = \sigma^2 \mathbb{I}_{N\times N}$$

y $\mathbb{A} = \begin{pmatrix} a_1 & \cdots & a_N \\ b_1 & \cdots & b_N \end{pmatrix}$ de dimensiones $2 \times N$. Por el resultado tenemos que

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_0 \end{pmatrix} = \mathbb{A}\mathbb{Y} \sim N_2\left(\mathbb{A}\mu_Y, \mathbb{A}\cdot\sigma^2\mathbb{I}\cdot\mathbb{A}'\right),$$

con

$$\mathbb{A}\mu_{\mathbb{Y}} = \begin{pmatrix} a_1(\beta_0+\beta_1 x_1) + \cdots + a_N(\beta_0+\beta_1 x_N) \\ b_1(\beta_0+\beta_1 x_1) + \cdots + a_N(\beta_0+\beta_1 x_N) \end{pmatrix}$$

$$= \begin{pmatrix} \beta_0 \sum_{i=1}^N a_i + \beta_1 \sum_{i=1}^N a_i x_i \\ \beta_0 \sum_{i=1}^N b_i + \beta_1 \sum_{i=1}^N b_i x_i \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_0 \end{pmatrix} \quad \text{ya que}$$

$$\sum_{i=1}^N a_i = 0 \; ; \; \sum_{i=1}^N a_i x_i = 1 \; ; \; \sum_{i=1}^N b_i = \sum_{i=1}^N \left(\tfrac{1}{N} - a_i\bar{x}\right) = 1 \; y \; \sum_{i=1}^N b_i x_i = \bar{x} - \bar{x} = 0$$

$$\mathbb{A}\cdot\sigma^2\mathbb{I}\,\mathbb{A}' = \sigma^2\mathbb{A}\mathbb{A}' = \sigma^2\begin{pmatrix} a_1, \dots, a_N \\ b_1, \dots, b_N \end{pmatrix}\begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_N & b_N \end{pmatrix}$$

$$= \sigma^2\begin{pmatrix} \sum_1^N a_i^2 & \sum_1^N a_i b_i \\ \sum_{i=1}^N b_i a_i & \sum_1^N b_i^2 \end{pmatrix}$$

$$= \sigma^2\begin{pmatrix} \dfrac{1}{\sum_{i=1}^N (x_i-\bar{x})^2} & -\dfrac{\bar{x}}{\sum_1^N (x_i-\bar{x})^2} \\ -\dfrac{\bar{x}}{\sum_1^N (x_i-\bar{x})^2} & \dfrac{\sum_1^N x_i^2}{N\sum_1^N (x_i-\bar{x})^2} \end{pmatrix}$$

como $\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_0 \end{pmatrix}$ es <u>conjuntamente</u> normal[1] y con matriz de varianzas covarianzas $\mathbb{A}(\sigma^2\mathbb{I})\mathbb{A}'$, tenemos que $\hat{\beta}_1$ y $\hat{\beta}_0$ son independientes, si y sólo si $\bar{x}=0$.

Por otra parte, como

$$\hat{\sigma}^{2\,(mv)} = \frac{1}{N}\sum_{i=1}^N \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2$$

(1) $\hat{\beta}_1 \sim N_1$ y $\hat{\beta}_0 \sim N_1$ con $cov(\hat{\beta}_1, \hat{\beta}_0) = 0 \not\Rightarrow$ independencia

se puede probar que

$$\frac{N}{\sigma^2}\hat{\sigma}^{2(MV)} \sim \chi^2_{(N-2)} \quad \cdots\cdots \quad (DV)$$

Usando este resultado tenemos que

$$\mathbb{E}\left(\frac{N}{\sigma^2}\hat{\sigma}^{2(MV)}\right) = N-2 \quad y \quad \mathbb{E}\left(\hat{\sigma}^{2(MV)}\right) = \left(\frac{N-2}{N}\right)\sigma^2$$

$$VAR\left(\frac{N}{\sigma^2}\hat{\sigma}^{2(MV)}\right) = 2(N-2) \quad y \quad VAR\left(\hat{\sigma}^{2(MV)}\right) = 2\frac{(N-2)}{N^2}\sigma^4$$

A partir de estos cálculos se puede definir un estimador insesgado para la varianza σ^2

$$\tilde{\sigma}^2 = \frac{1}{N-2}\sum_{i=1}^{N}\left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i\right)^2$$

PROPOSICION: $\hat{\beta}_0$ y $\hat{\beta}_1$ son independientes de $\hat{\sigma}^2$

Dem

$$D_j \equiv Y_j - \hat{\beta}_0 - \hat{\beta}_1 x_i \quad j=1,\dots,N$$

cada D_j es una combinacion lineal de las v.a. Normales $Y_1, Y_2, \dots, Y_N$, por lo que D_j es es una v.a. Normal $\forall j = 1, \dots, N$. De hecho usando la proposición ★ en la página 68 el vector aleatorio $\mathbb{D}_j = (D_j, \hat{\beta}_1, \hat{\beta}_0)$; $j=1,\dots,N$, tiene distribución Normal Multivariada. En estas condiciones, para verificar la independencia entre los coeficientes $\hat{\beta}_0$, $\hat{\beta}_1$ y $\hat{\sigma}^2$ basta comprobar que

$$\operatorname{cov}(D_j, \hat{\beta}_1) = 0 \quad y \quad \operatorname{cov}(D_j, \hat{\beta}_0) = 0 \quad \dots \quad (A)$$

$$\forall j = 1, 2, \dots, N.$$

Si sucede lo que indican las relaciones en (A), entonces $\hat{\beta}_1$ es independiente de cada D_j $\forall j=1,\dots,N$ luego $\hat{\beta}_1$ es independiente de D_j^2 $\forall j=1,2,\dots,N$ y por tanto $\hat{\beta}_1$ es independiente de $\frac{1}{N}\sum_{j=1}^{N} D_j^2 = \hat{\sigma}^2$. El argumento es análogo para $\hat{\beta}_0$.