

Teoría detrás de las herramientas de SageMath

Miguel Calvo Arnal, Juan Guerrero Marcos y Andrés Laín Sanclemente

Índice

1. Distribución normal o gaussiana	3
1.1. Distribución χ^2_ν	9
2. Propagación de errores	10
3. Ajustes lineales con errores únicamente en las ordenadas	12
4. Ajustes a funciones arbitrarias con errores únicamente en las ordenadas	18

Notación 0.1. Para las medidas no usamos ningún tipo de símbolo en especial x , para la variable aleatoria asociada a una medida x usaremos la notación \dot{x} , para la variable asociada a la variable aleatoria \dot{x} y que aparece como argumento en su función de densidad de probabilidad y de distribución usaremos \dot{x} . Para un valor ideal o «real» de una medida x usaremos \tilde{x} y, por último, para denotar un estimador de una variable aleatoria \dot{x} usaremos \hat{x} .

1. Distribución normal o gaussiana

Proposición 1.1. Sea \dot{x} una variable aleatoria gaussiana de centro μ y varianza σ^2 . Entonces $\dot{y} = \dot{x} + b$, donde $b \in \mathbb{R}$ es una constante, es también una variable aleatoria gaussiana de varianza σ^2 , pero de centro $\mu + b$.

$$\dot{x} \sim \mathcal{N}(\mu, \sigma^2) \implies \dot{y} = \dot{x} + b \sim \mathcal{N}(\mu + b, \sigma^2) \quad \forall b \in \mathbb{R}$$

Demostración. La función de distribución de la variable aleatoria \dot{y} se relaciona con la de \dot{x} mediante:

$$F_{\dot{y}}(\dot{y}) = P(\dot{y} < \dot{y}) = P(\dot{x} + b < \dot{y}) = P(\dot{x} < \dot{y} - b) = F_{\dot{x}}(\dot{y} - b)$$

Derivando la expresión anterior, obtenemos:

$$\begin{aligned} f_{\dot{y}}(\dot{y}) &= \frac{dF_{\dot{y}}}{d\dot{y}}(\dot{y}) = \frac{dF_{\dot{x}}}{d\dot{x}}(\dot{y} - b) = f_{\dot{x}}(\dot{y} - b) = \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\dot{y} - b - \mu)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\dot{y} - (\mu + b))^2}{2\sigma^2}\right) \end{aligned}$$

que claramente se corresponde con la la función de densidad de una variable aleatoria gaussiana de centro $\mu + b$ y varianza σ^2 . Q.E.D.

Proposición 1.2. Sea \dot{x} una variable aleatoria gaussiana de centro μ y varianza σ^2 . Entonces $\dot{y} = a\dot{x}$, donde $a \in \mathbb{R}$ es una constante, es también una variable aleatoria normal, pero de centro $a\mu$ y de varianza $a^2\sigma^2$.

$$\dot{x} \sim \mathcal{N}(\mu, \sigma^2) \implies \dot{y} = a\dot{x} \sim \mathcal{N}(a\mu, a^2\sigma^2) \quad \forall a \in \mathbb{R}$$

Demostración. La función de densidad de la variable aleatoria \dot{y} se relaciona con la de \dot{x} mediante:

$$F_{\dot{y}}(\dot{y}) = P(\dot{y} < \dot{y}) = P(a\dot{x} < \dot{y}) = P\left(\dot{x} < \frac{\dot{y}}{a}\right) = F_{\dot{x}}\left(\frac{\dot{y}}{a}\right)$$

Derivando la expresión anterior, obtenemos:

$$\begin{aligned} f_{\dot{y}}(\dot{y}) &= \frac{dF_{\dot{y}}}{d\dot{y}}(\dot{y}) = \frac{dF_{\dot{x}}}{d\dot{x}}\left(\frac{\dot{y}}{a}\right) \frac{1}{a} = \frac{1}{a} f_{\dot{x}}\left(\frac{\dot{y}}{a}\right) = \\ &= \frac{1}{a} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(\frac{\dot{y}}{a} - \mu\right)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi a^2\sigma^2}} \exp\left(-\frac{\left(\frac{\dot{y} - a\mu}{a}\right)^2}{2\sigma^2}\right) = \\ &= \frac{1}{\sqrt{2\pi a^2\sigma^2}} \exp\left(-\frac{(\dot{y} - a\mu)^2}{2a^2\sigma^2}\right) \end{aligned}$$

Y lo anterior es claramente la función de densidad de probabilidad de una variable aleatoria gaussiana de centro $a\mu$ y varianza $a^2\sigma^2$. Q.E.D.

Proposición 1.3. *Toda variable aleatoria normal puede «normalizarse» a una variable aleatoria gaussiana de centro 0 y varianza 1.*

$$\hat{x} \sim \mathcal{N}(\mu, \sigma^2) \implies \hat{z} = \frac{\hat{x} - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

Demostración. Sea $\hat{x} \sim \mathcal{N}(\mu, \sigma^2)$. Por la proposición 1.1 en la página anterior, la variable aleatoria $\hat{y} = \hat{x} - \mu$, cumple $\hat{y} \sim \mathcal{N}(0, \sigma^2)$. Por último, por la proposición 1.2 en la página anterior, la variable aleatoria $\hat{z} = \frac{1}{\sigma}\hat{y}$ satisface $\hat{z} \sim \mathcal{N}(0, 1)$. Por tanto:

$$\hat{z} = \frac{\hat{y}}{\sigma} = \frac{\hat{x} - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

Q.E.D.

Proposición 1.4. *La combinación lineal de dos variables aleatorias gaussianas normalizadas es otra variable aleatoria normal de centro 0 y varianza $a^2 + b^2$, siendo a y b los coeficientes de la combinación lineal.*

$$\hat{x}_1, \hat{x}_2 \sim \mathcal{N}(0, 1) \implies \hat{y} = a\hat{x}_1 + b\hat{x}_2 \sim \mathcal{N}(0, a^2 + b^2)$$

con $a, b \in \mathbb{R}$.

Demostración. Intentemos obtener la función de distribución de la variable \hat{y} :

$$\begin{aligned} F_{\hat{y}}(\hat{y}) &= \mathbb{P}(\hat{y} < \hat{y}) = \mathbb{P}(a\hat{x}_1 + b\hat{x}_2 < \hat{y}) = \iint_{\{a\hat{x}_1 + b\hat{x}_2 < \hat{y}\}} f_{\hat{x}_1}(\hat{x}_1) f_{\hat{x}_2}(\hat{x}_2) d\hat{x}_1 d\hat{x}_2 = \\ &= \iint_{\{a\hat{x}_1 + b\hat{x}_2 < \hat{y}\}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\hat{x}_1^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\hat{x}_2^2}{2}\right) d\hat{x}_1 d\hat{x}_2 = \\ &= \iint_{\{a\hat{x}_1 + b\hat{x}_2 < \hat{y}\}} \frac{1}{2\pi} \exp\left(-\frac{\hat{x}_1^2}{2} - \frac{\hat{x}_2^2}{2}\right) d\hat{x}_1 d\hat{x}_2 = \\ &= \iint_{\{a\hat{x}_1 + b\hat{x}_2 < \hat{y}\}} \frac{1}{2\pi} \exp\left(-\frac{\hat{x}_1^2 + \hat{x}_2^2}{2}\right) d\hat{x}_1 d\hat{x}_2 \end{aligned}$$

A continuación, hacemos el cambio de variable:

$$\dot{u}_1 = a\dot{x}_1 + b\dot{x}_2$$

$$\dot{u}_2 = a\dot{x}_1 - b\dot{x}_2$$

Sumando las dos ecuaciones anteriores, podemos despejar \dot{x}_1 :

$$\dot{u}_1 + \dot{u}_2 = 2a\dot{x}_1 \iff \dot{x}_1 = \frac{\dot{u}_1 + \dot{u}_2}{2a}$$

Y restando las dos ecuaciones anteriores, podemos despejar \dot{x}_2 :

$$\dot{u}_1 - \dot{u}_2 = 2b\dot{x}_2 \iff \dot{x}_2 = \frac{\dot{u}_1 - \dot{u}_2}{2b}$$

Calculamos ahora la matriz jacobiana del cambio $(\dot{u}_1, \dot{u}_2) \longrightarrow (\dot{x}_1, \dot{x}_2)$:

$$J = \begin{pmatrix} \frac{\partial \dot{x}_1}{\partial \dot{u}_1} & \frac{\partial \dot{x}_1}{\partial \dot{u}_2} \\ \frac{\partial \dot{x}_2}{\partial \dot{u}_1} & \frac{\partial \dot{x}_2}{\partial \dot{u}_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2a} & \frac{1}{2a} \\ \frac{1}{2b} & -\frac{1}{2b} \end{pmatrix}$$

$$|\det J| = \left| -\frac{1}{4ab} - \frac{1}{4ab} \right| = \frac{1}{2ab}$$

Por otra parte, el conjunto de integración queda:

$$\{(\dot{x}_1, \dot{x}_2) \in \mathbb{R}^2 \text{ t.q. } a\dot{x}_1 + b\dot{x}_2 < \dot{y}\} = \{(\dot{u}_1, \dot{u}_2) \in (-\infty, \dot{y}) \times \mathbb{R}\}$$

Con este cambio de variable el dominio de integración ha quedado infinitamente más sencillo de manejar. De esta forma, por el teorema del cambio de variable, la integral resulta:

$$\begin{aligned} F_{\dot{y}}(\dot{y}) &= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{2ab} \frac{1}{2\pi} \exp\left(-\frac{\left(\frac{\dot{u}_1+\dot{u}_2}{2a}\right)^2 + \left(\frac{\dot{u}_1-\dot{u}_2}{2b}\right)^2}{2}\right) d\dot{u}_1 d\dot{u}_2 = \\ &= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left(-\frac{\frac{\dot{u}_1^2+\dot{u}_2^2+2\dot{u}_1\dot{u}_2}{4a^2} + \frac{\dot{u}_1^2+\dot{u}_2^2-2\dot{u}_1\dot{u}_2}{4b^2}}{2}\right) d\dot{u}_1 d\dot{u}_2 = \\ &= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left(-\frac{\frac{\dot{u}_1^2+\dot{u}_2^2+2\dot{u}_1\dot{u}_2}{a^2} + \frac{\dot{u}_1^2+\dot{u}_2^2-2\dot{u}_1\dot{u}_2}{b^2}}{8}\right) d\dot{u}_1 d\dot{u}_2 = \\ &= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left[-\frac{b^2(\dot{u}_1^2 + \dot{u}_2^2 + 2\dot{u}_1\dot{u}_2) + a^2(\dot{u}_1^2 + \dot{u}_2^2 - 2\dot{u}_1\dot{u}_2)}{8a^2b^2}\right] d\dot{u}_1 d\dot{u}_2 = \\ &= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp\left[-\frac{(b^2 + a^2)\dot{u}_1^2 + (b^2 + a^2)\dot{u}_2^2 + 2\dot{u}_1\dot{u}_2(b^2 - a^2)}{8a^2b^2}\right] d\dot{u}_1 d\dot{u}_2 \quad (1.1) \end{aligned}$$

A continuación, nuestro propósito es completar cuadrados con el fin de poder hacer la integral en la variable \dot{u}_2 . Para ello, notemos:

$$\begin{aligned} &\left(\sqrt{b^2 + a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2 - a^2)}{\sqrt{b^2 + a^2}}\right)^2 = \\ &= (b^2 + a^2)\dot{u}_2^2 + \dot{u}_1^2 \frac{(b^2 - a^2)^2}{b^2 + a^2} + 2\dot{u}_1\dot{u}_2(b^2 - a^2) \iff \\ &\iff \left(\sqrt{b^2 + a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2 - a^2)}{\sqrt{b^2 + a^2}}\right)^2 - \dot{u}_1^2 \frac{(b^2 - a^2)^2}{b^2 + a^2} = (b^2 + a^2)\dot{u}_2^2 + 2\dot{u}_1\dot{u}_2(b^2 - a^2) \end{aligned}$$

De esta forma, podemos expresar el exponente del integrando de la expresión 1.1 como:

$$\begin{aligned} &-\frac{\left(\sqrt{b^2 + a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2 - a^2)}{\sqrt{b^2 + a^2}}\right)^2 - \dot{u}_1^2 \frac{(b^2 - a^2)^2}{b^2 + a^2} + (b^2 + a^2)\dot{u}_1^2}{8a^2b^2} = \\ &= -\frac{\left(\sqrt{b^2 + a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2 - a^2)}{\sqrt{b^2 + a^2}}\right)^2 + \frac{(b^2 + a^2)^2 - (b^2 - a^2)^2}{b^2 + a^2}\dot{u}_1^2}{8a^2b^2} = \\ &= -\frac{\left(\sqrt{b^2 + a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2 - a^2)}{\sqrt{b^2 + a^2}}\right)^2 + \frac{b^4 + a^4 + 2a^2b^2 - b^4 - a^4 + 2a^2b^2}{b^2 + a^2}\dot{u}_1^2}{8a^2b^2} = \\ &= -\frac{\left(\sqrt{b^2 + a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2 - a^2)}{\sqrt{b^2 + a^2}}\right)^2 + \frac{4a^2b^2}{b^2 + a^2}\dot{u}_1^2}{8a^2b^2} = \\ &= -\frac{1}{2} \left(\frac{\sqrt{b^2 + a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2 - a^2)}{\sqrt{b^2 + a^2}}}{2ab}\right)^2 - \frac{\dot{u}_1^2}{2(a^2 + b^2)} \end{aligned}$$

Sustituyendo de vuelta en la integral de la expresión 1.1 en la página anterior, se tiene:

$$\begin{aligned} F_{\dot{y}}(\dot{y}) &= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp \left[-\frac{1}{2} \left(\frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{2ab} \right)^2 - \frac{\dot{u}_1^2}{2(a^2+b^2)} \right] d\dot{u}_1 d\dot{u}_2 = \\ &= \int_{\dot{u}_1=-\infty}^{\dot{y}} \int_{\dot{u}_2=-\infty}^{\infty} \frac{1}{4\pi ab} \exp \left[-\frac{1}{2} \left(\frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{2ab} \right)^2 \right] \exp \left[-\frac{\dot{u}_1^2}{2(a^2+b^2)} \right] d\dot{u}_1 d\dot{u}_2 \end{aligned}$$

Por el teorema de Fubini, se tiene:

$$F_{\dot{y}}(\dot{y}) = \int_{\dot{u}_1=-\infty}^{\dot{y}} \frac{1}{4\pi ab} \exp \left[-\frac{\dot{u}_1^2}{2(a^2+b^2)} \right] \underbrace{\left(\int_{\dot{u}_2=-\infty}^{\infty} \exp \left[-\frac{1}{2} \left(\frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{2ab} \right)^2 \right] d\dot{u}_2 \right)}_{=:I} d\dot{u}_1 \quad (1.2)$$

Para proseguir, solucionaremos la integral I de forma separada. Para ello, recuérdese que dentro de esta integral \dot{u}_1 no es más que una constante. De esta forma, hacemos el cambio de variable:

$$\begin{aligned} \dot{t} &:= \frac{\sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{2ab} \iff 2ab\dot{t} = \sqrt{b^2+a^2}\dot{u}_2 + \dot{u}_1 \frac{(b^2-a^2)}{\sqrt{b^2+a^2}} \iff \\ &\iff 2ab\dot{t} - \dot{u}_1 \frac{(b^2-a^2)}{\sqrt{b^2+a^2}} = \sqrt{b^2+a^2}\dot{u}_2 \iff \\ &\iff \dot{u}_2 = \frac{2ab\dot{t} - \dot{u}_1 \frac{(b^2-a^2)}{\sqrt{b^2+a^2}}}{\sqrt{a^2+b^2}} \end{aligned}$$

Derivando, se tiene:

$$d\dot{u}_2 = \frac{2ab}{\sqrt{a^2+b^2}} d\dot{t}$$

Además:

$$\begin{aligned} \dot{u}_2 \rightarrow \infty &\iff \dot{t} \rightarrow \infty \\ \dot{u}_2 \rightarrow -\infty &\iff \dot{t} \rightarrow -\infty \end{aligned}$$

Por consiguiente, la integral I queda:

$$\begin{aligned} I &= \int_{\dot{t}=-\infty}^{\infty} \exp \left(-\frac{1}{2}\dot{t}^2 \right) \frac{2ab}{\sqrt{a^2+b^2}} d\dot{t} = \\ &= \frac{2ab}{\sqrt{a^2+b^2}} \sqrt{2\pi} \underbrace{\int_{\dot{t}=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2}\dot{t}^2 \right) d\dot{t}}_{*} \end{aligned}$$

donde el término marcado es claramente la función de densidad de una variable aleatoria gaussiana de centro 0 y varianza unidad: $\dot{t} \sim \mathcal{N}(0, 1)$. Por tanto, la integral que nos resta por evaluar debe ser la unidad por la definición de función de densidad. Así, se llega a:

$$I = \frac{2ab}{\sqrt{a^2+b^2}} \sqrt{2\pi}$$

Sustituyendo de vuelta en la integral 1.2, se tiene:

$$F_{\dot{y}}(\dot{y}) = \int_{\dot{u}_1=-\infty}^{\dot{y}} \frac{1}{4\pi ab} \exp \left(-\frac{\dot{u}_1^2}{2(a^2+b^2)} \right) \frac{2ab}{\sqrt{a^2+b^2}} \sqrt{2\pi} d\dot{u}_1 =$$

$$\begin{aligned}
&= \int_{\dot{u}_1=-\infty}^{\dot{y}} \frac{\sqrt{2\pi}}{2\pi} \frac{1}{\sqrt{a^2 + b^2}} \exp\left(-\frac{\dot{u}_1^2}{2(a^2 + b^2)}\right) d\dot{u}_1 = \\
&= \int_{\dot{u}_1=-\infty}^{\dot{y}} \frac{1}{\sqrt{2\pi(a^2 + b^2)}} \exp\left(-\frac{\dot{u}_1^2}{2(a^2 + b^2)}\right) d\dot{u}_1
\end{aligned}$$

A continuación, por el Teorema Fundamental del Cálculo Integral, se tiene:

$$f_{\dot{y}}(\dot{y}) = \frac{dF_{\dot{y}}}{d\dot{y}}(\dot{y}) = \frac{1}{\sqrt{2\pi(a^2 + b^2)}} \exp\left(-\frac{\dot{y}^2}{2(a^2 + b^2)}\right)$$

que es claramente la función de densidad de una variable aleatoria normal de centro 0 y de varianza $a^2 + b^2$. Por ende:

$$\dot{y} \sim \mathcal{N}(0, a^2 + b^2)$$

Q.E.D.

Corolario 1.1. *La combinación lineal de dos variables aleatorias gaussianas es otra variable aleatoria gaussiana. En concreto:*

$$\left. \begin{aligned} \dot{x}_1 &\sim \mathcal{N}(\mu_1, \sigma_1^2) \\ \dot{x}_2 &\sim \mathcal{N}(\mu_2, \sigma_2^2) \end{aligned} \right\} \implies \dot{y} = a\dot{x}_1 + b\dot{x}_2 \sim \mathcal{N}(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$$

con $a, b \in \mathbb{R}$.

Demostración. Primero, nótese que:

$$\begin{aligned}
\dot{y} &= a\dot{x}_1 + b\dot{x}_2 = a(\dot{x}_1 - \mu_1 + \mu_1) + b(\dot{x}_2 - \mu_2 + \mu_2) = \\
&= a(\dot{x}_1 - \mu_1) + a\mu_1 + b(\dot{x}_2 - \mu_2) + b\mu_2 = \\
&= a\sigma_1 \left(\frac{\dot{x}_1 - \mu_1}{\sigma_1} \right) + a\mu_1 + b\sigma_2 \left(\frac{\dot{x}_2 - \mu_2}{\sigma_2} \right) + b\mu_2 = \\
&= a\mu_1 + b\mu_2 + a\sigma_1 \left(\frac{\dot{x}_1 - \mu_1}{\sigma_1} \right) + b\sigma_2 \left(\frac{\dot{x}_2 - \mu_2}{\sigma_2} \right)
\end{aligned} \tag{1.3}$$

Ahora bien, por la proposición 1.3 en la página 4:

$$\dot{z}_1 := \frac{\dot{x}_1 - \mu_1}{\sigma_1} \quad \dot{z}_2 := \frac{\dot{x}_2 - \mu_2}{\sigma_2}$$

son ambas variables aleatorias normales estándar: $\dot{z}_1, \dot{z}_2 \sim \mathcal{N}(0, 1)$. Por tanto, el término;

$$a\sigma_1 \left(\frac{\dot{x}_1 - \mu_1}{\sigma_1} \right) + b\sigma_2 \left(\frac{\dot{x}_2 - \mu_2}{\sigma_2} \right) = a\sigma_1 \dot{z}_1 + b\sigma_2 \dot{z}_2$$

es una combinación lineal de variables aleatorias gaussianas normalizadas. Por la proposición 1.4 en la página 4, se tiene:

$$\dot{Z} := a\sigma_1 \left(\frac{\dot{x}_1 - \mu_1}{\sigma_1} \right) + b\sigma_2 \left(\frac{\dot{x}_2 - \mu_2}{\sigma_2} \right) \sim \mathcal{N}(0, a^2\sigma_1^2 + b^2\sigma_2^2)$$

Sustituyendo de nuevo en la ecuación 1.3, obtenemos:

$$\dot{y} = a\mu_1 + b\mu_2 + \dot{Z}$$

donde $\dot{Z} \sim \mathcal{N}(0, a^2\sigma_1^2 + b^2\sigma_2^2)$. Por la proposición 1.1 en la página 3, se tiene que:

$$\dot{y} \sim \mathcal{N}(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$$

Q.E.D.

Corolario 1.2 (¡Importante!). *Cualquier combinación lineal de variables aleatorias gaussianas es otra variable aleatoria gaussiana.*

$$\dot{x}_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad \forall i = 1, \dots, n \implies \dot{y} = \sum_{i=1}^n \alpha_i \dot{x}_i \sim \mathcal{N}\left(\sum_{i=1}^n \alpha_i \mu_i, \sum_{i=1}^n \alpha_i^2 \sigma_i^2\right)$$

donde $\vec{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$.

Demostración. Probaremos el resultado por inducción sobre n . Por el corolario 1.1 en la página anterior, el resultado se cumple para $n = 2$. Supongamos que se cumple para n y veamos si se cumple para $n + 1$. Es decir, sean $\{\dot{x}_i\}_{i=1}^{n+1}$ variables aleatorias tales que:

$$\dot{x}_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad \forall i = 1, \dots, n$$

y sea $\alpha \in \mathbb{R}^{n+1}$. Entonces:

$$\dot{y} = \sum_{i=1}^{n+1} \alpha_i \dot{x}_i = \sum_{i=1}^n \alpha_i \dot{x}_i + \alpha_{n+1} \dot{x}_{n+1}$$

Por hipótesis de inducción:

$$\dot{z} := \sum_{i=1}^n \alpha_i \dot{x}_i \sim \mathcal{N}\left(\sum_{i=1}^n \alpha_i \mu_i, \sum_{i=1}^n \alpha_i^2 \sigma_i^2\right)$$

Entonces, nos queda por evaluar la suma de dos variables gaussianas:

$$\dot{y} = \dot{z} + \alpha_{n+1} \dot{x}_{n+1}$$

Por el corolario 1.1 en la página anterior, se tiene:

$$\begin{aligned} \dot{y} &\sim \mathcal{N}\left(\sum_{i=1}^n \alpha_i \mu_i + \alpha_{n+1} \dot{x}_{n+1}, \sum_{i=1}^n \alpha_i^2 \sigma_i^2 + \alpha_{n+1}^2 \sigma_{n+1}^2\right) = \\ &= \mathcal{N}\left(\sum_{i=1}^{n+1} \alpha_i \mu_i, \sum_{i=1}^{n+1} \alpha_i^2 \sigma_i^2\right) \end{aligned}$$

por lo que el resultado se cumple para $n + 1$, lo que completa la demostración. Q.E.D.

Corolario 1.3. *Toda combinación lineal normalizada de variables gaussianas normalizadas es otra variable aleatoria gaussiana normalizada.*

$$\left. \begin{array}{l} \dot{x}_i \sim \mathcal{N}(0, 1) \quad \forall i = 1, \dots, n \\ \vec{\alpha} = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n \text{ t.q. } \|\vec{\alpha}\| = 1 \end{array} \right\} \implies \alpha^T \vec{\dot{x}} = \sum_{i=1}^n \alpha_i \dot{x}_i \sim \mathcal{N}(0, 1)$$

Demostración. Por el corolario 1.1 en la página anterior, se tiene que:

$$\begin{aligned} \sum_{i=1}^n \alpha_i \dot{x}_i &\sim \mathcal{N}\left(\sum_{i=1}^n \alpha_i \cdot 0, \sum_{i=1}^n \alpha_i^2 \cdot 1^2\right) = \mathcal{N}\left(0, \sum_{i=1}^n \alpha_i^2\right) = \\ &= \mathcal{N}\left(0, \|\vec{\alpha}\|^2\right) = \mathcal{N}(0, 1^2) = \mathcal{N}(0, 1) \end{aligned}$$

dado que $\vec{\alpha} \in \mathbb{R}^n$ era de norma unidad. Q.E.D.

1.1. Distribución χ^2_ν

Proposición 1.5. *Dado un vector de variables aleatorias gaussianas normalizadas $\vec{x} = (\dot{x}_1, \dots, \dot{x}_n)$ que pertenece a un espacio vectorial de dimensión $n - k$, entonces, el cuadrado de la norma euclídea del vector sigue una distribución χ^2 con $n - k$ grados de libertad. Es decir:*

$$\left. \begin{array}{l} \dot{x}_i \sim \mathcal{N}(0, 1) \quad \forall i = 1, \dots, n \\ \dim(\langle \vec{x} \rangle) = n - k \end{array} \right\} \implies \|\vec{x}\|^2 = \dot{x}_1^2 + \dots + \dot{x}_n^2 = \sum_{i=1}^n \dot{x}_i^2 \sim \chi_{n-k}^2$$

Demostración. Llamemos W al espacio vectorial generado por el vector aleatorio \vec{x} , que tiene, por hipótesis, dimensión $n - k$. Entonces, sabemos que podemos descomponer \mathbb{R}^n como suma directa de W y de su complementario ortogonal. Es decir:

$$\mathbb{R}^n = W \oplus W^\perp \quad (1.4)$$

A continuación, sea $\{\vec{a}_1, \dots, \vec{a}_{n-k}\}$ una base ortonormal de W (que sabemos que siempre existe al menos una) y sea $\{\vec{b}_1, \dots, \vec{b}_k\}$ una base ortonormal de W^\perp . Entonces, por la ecuación 1.4 $U := \{\vec{a}_1, \dots, \vec{a}_{n-k}\} \cup \{\vec{b}_1, \dots, \vec{b}_k\} = \{\vec{a}_1, \dots, \vec{a}_{n-k}, \vec{b}_1, \dots, \vec{b}_k\}$ es una base ortonormal de \mathbb{R}^n .

Ahora bien, en este momento, estamos expresando el vector \vec{x} en base canónica (E), que es una base ortonormal. Por tanto, el cambio de base entre E y U vendrá dado por una matriz ortogonal. Es decir, existe $Q \in \mathbb{R}^{n \times n}$ tal que $Q^T Q = \mathbb{I}$ y $\vec{z} \in \mathbb{R}^n$ tal que:

$$\vec{z} = Q \vec{x} \quad (1.5)$$

es la representación de \vec{x} en base U . Además, como Q es una matriz ortogonal, sus columnas son vectores de norma unidad de \mathbb{R}^n . De esta forma, cada \dot{z}_i es una combinación lineal normalizada de las $\{\dot{x}_i\}_{i=1}^n$ y, por el lema 1.3 en la página anterior, cada \dot{z}_i es también una variable aleatoria gaussiana normalizada. Como $\vec{x} \in W$, sabemos que las últimas k componentes de \vec{z} serán, necesariamente, nulas. Es decir:

$$\dot{z}_{n-k+1} = \dot{z}_{n-k+2} = \dots = \dot{z}_n = 0 \quad (1.6)$$

Por otra parte:

$$\|\vec{x}\|^2 = \sum_{i=1}^n \dot{x}_i^2 = \vec{x}^T \vec{x} = \vec{x}^T \mathbb{I} \vec{x} = \vec{x}^T Q^T Q \vec{x}$$

ya que $Q^T Q = \mathbb{I}$. Si seguimos operando, se tiene:

$$\|\vec{x}\|^2 = (Q \vec{x})^T Q \vec{x} = \vec{z}^T \vec{z} = \sum_{i=1}^n \dot{z}_i^2$$

ya que era $\vec{z} = Q \vec{x}$ por la ecuación 1.5. Ahora bien, aplicando la ecuación 1.6, se obtiene:

$$\|\vec{x}\|_2^2 = \sum_{i=1}^{n-k} \dot{z}_i^2 + \sum_{i=n-k+1}^n \dot{z}_i^2 = \sum_{i=1}^{n-k} \dot{z}_i^2$$

Como $\{\dot{z}_i\}_{i=1}^{n-k}$ es un conjunto de $n - k$ variables aleatorias gaussianas normalizadas independientes entre sí, entonces:

$$\|\vec{x}\|_2^2 = \sum_{i=1}^n \dot{x}_i^2 = \sum_{i=1}^{n-k} \dot{z}_i^2 \sim \chi_{n-k}^2$$

por definición de la distribución χ_{n-k}^2 .

Q.E.D.

2. Propagación de errores

Teorema 2.1. Sea Ω un conjunto abierto de \mathbb{R}^n . A continuación, sea $\vec{x} = (\dot{x}_1, \dots, \dot{x}_n)$ un vector de variables aleatorias y sea $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una función escalar de clase $C^{(1)}$. Al primer orden no nulo, se tiene:

$$\begin{aligned} \mathbb{E} \left(f \left(\vec{x} \right) \right) &= f \left(\mathbb{E} \left(\vec{x} \right) \right) + o(2) \\ \text{Var} \left(f \left(\vec{x} \right) \right) &= \left[\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \right]^T \text{Cov} \left(\vec{x} \right) \vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) + o(3) \end{aligned}$$

Demostración. Como la función f es de clase $C^{(1)}$, admite un desarrollo de Taylor de orden 1 en torno a $\mathbb{E} \left(\vec{x} \right)$.

$$f \left(\vec{x} \right) = f \left(\mathbb{E} \left(\vec{x} \right) \right) + \vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \left[\vec{x} - \mathbb{E} \left(\vec{x} \right) \right] + o(2) \quad (2.1)$$

donde es importante distinguir $\dot{\vec{x}}$ de \vec{x} . A continuación, obtengamos el valor esperado de $f \left(\vec{x} \right)$:

$$\begin{aligned} \mathbb{E} \left[f \left(\vec{x} \right) \right] &= \mathbb{E} \left[f \left(\mathbb{E} \left(\vec{x} \right) \right) + \vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \left[\vec{x} - \mathbb{E} \left(\vec{x} \right) \right] + o(2) \right] = \\ &= f \left(\mathbb{E} \left(\vec{x} \right) \right) + \vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \mathbb{E} \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) + o(2) \end{aligned}$$

Ahora bien, recordemos que $\mathbb{E} \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) = 0$ para cualquier variable aleatoria. De esta forma, tenemos:

$$\mathbb{E} \left[f \left(\vec{x} \right) \right] = f \left(\mathbb{E} \left(\vec{x} \right) \right) + o(2)$$

con lo que llegamos a la primera parte del enunciado.

Vamos con la varianza. Aplicaremos la expresión:

$$\text{Var} (\dot{y}) = \mathbb{E} (\dot{y}^2) - (\mathbb{E} (\dot{y}))^2$$

a la variable aleatoria $\dot{y} = f \left(\vec{x} \right)$, que sabemos que debe cumplirse para cualquier variable aleatoria. Elevando al cuadrado la expresión 2.1, se tiene:

$$f \left(\vec{x} \right)^2 = f \left(\mathbb{E} \left(\vec{x} \right) \right)^2 + \left[\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) \right]^2 + 2f \left(\mathbb{E} \left(\vec{x} \right) \right) \vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) + o(3)$$

agrupando en $o(3)$ todos los términos de orden 3 y superior. A continuación, podemos calcular el valor esperado de $\left[f \left(\vec{x} \right) \right]^2$:

$$\begin{aligned} \mathbb{E} \left[f \left(\vec{x} \right)^2 \right] &= \mathbb{E} \left[f \left(\mathbb{E} \left(\vec{x} \right) \right)^2 + \left[\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) \right]^2 + \right. \\ &\quad \left. + 2f \left(\mathbb{E} \left(\vec{x} \right) \right) \vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) + o(3) \right] = \\ &= f \left(\mathbb{E} \left(\vec{x} \right) \right)^2 + \mathbb{E} \left[\left(\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) \right)^2 \right] + 2f \left(\mathbb{E} \left(\vec{x} \right) \right) \vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \mathbb{E} \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) + o(3) \end{aligned}$$

De nuevo, como es $\mathbb{E} \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) = 0$, se tiene:

$$\mathbb{E} \left[f \left(\vec{x} \right)^2 \right] = f \left(\mathbb{E} \left(\vec{x} \right) \right)^2 + \mathbb{E} \left[\left(\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \cdot \left(\vec{x} - \mathbb{E} \left(\vec{x} \right) \right) \right)^2 \right] + o(3)$$

De esta forma, ya estamos en disposición de obtener la varianza de $f\left(\overset{\circ}{\vec{x}}\right)$:

$$\begin{aligned}\text{Var}\left(f\left(\overset{\circ}{\vec{x}}\right)\right) &= \text{E}\left(f\left(\overset{\circ}{\vec{x}}\right)^2\right) - \left[\text{E}\left(f\left(\overset{\circ}{\vec{x}}\right)\right)\right]^2 = \\ &= f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)^2 + \text{E}\left[\left(\vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) \cdot \left(\overset{\circ}{\vec{x}} - \text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right)^2\right] - f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)^2 + o(3) = \\ &= \text{E}\left[\left(\vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) \cdot \left(\overset{\circ}{\vec{x}} - \text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right)^2\right] + o(3)\end{aligned}$$

Para proseguir, desarrollamos el producto escalar dentro del cuadrado:

$$\text{Var}\left(f\left(\overset{\circ}{\vec{x}}\right)\right) = \text{E}\left[\left(\sum_{i=1}^n \frac{\partial f}{\partial \dot{x}_i}\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) (\dot{x}_i - \text{E}(\dot{x}_i))\right)^2\right] + o(3)$$

Ahora, expandimos el cuadrado anterior como el doble sumatorio de cualquier combinación de términos:

$$\begin{aligned}\text{Var}\left(f\left(\overset{\circ}{\vec{x}}\right)\right) &= \text{E}\left[\sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial \dot{x}_i}\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) \frac{\partial f}{\partial \dot{x}_j}\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) (\dot{x}_i - \text{E}(\dot{x}_i)) (\dot{x}_j - \text{E}(\dot{x}_j))\right] + o(3) = \\ &= \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial \dot{x}_i}\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) \frac{\partial f}{\partial \dot{x}_j}\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) \text{E}[(\dot{x}_i - \text{E}(\dot{x}_i)) (\dot{x}_j - \text{E}(\dot{x}_j))] + o(3)\end{aligned}$$

Justo el término $\text{E}[(\dot{x}_i - \text{E}(\dot{x}_i)) (\dot{x}_j - \text{E}(\dot{x}_j))]$ es la definición de covarianza entre las variables aleatorias \dot{x}_i y \dot{x}_j . De esta forma, obtenemos:

$$\text{Var}\left(f\left(\overset{\circ}{\vec{x}}\right)\right) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial f}{\partial \dot{x}_i}\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) \frac{\partial f}{\partial \dot{x}_j}\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) \text{Cov}(\dot{x}_i, \dot{x}_j) + o(3)$$

A continuación, es nuestra intención convertir lo anterior en un producto vector*matriz*vector. Para ello, recordamos:

$$\begin{aligned}\vec{\nabla}f &= \left(\frac{\partial f}{\partial \dot{x}_1}, \dots, \frac{\partial f}{\partial \dot{x}_n}\right) \\ \text{Cov}\left(\overset{\circ}{\vec{x}}\right) &= \begin{pmatrix} \text{Var}(\dot{x}_1) & \text{Cov}(\dot{x}_1, \dot{x}_2) & \cdots & \text{Cov}(\dot{x}_1, \dot{x}_{n-1}) & \text{Cov}(\dot{x}_1, \dot{x}_n) \\ \text{Cov}(\dot{x}_2, \dot{x}_1) & \ddots & & & \text{Cov}(\dot{x}_2, \dot{x}_n) \\ \vdots & & \ddots & & \vdots \\ \text{Cov}(\dot{x}_{n-1}, \dot{x}_1) & & & \ddots & \text{Cov}(\dot{x}_{n-1}, \dot{x}_n) \\ \text{Cov}(\dot{x}_n, \dot{x}_1) & \text{Cov}(\dot{x}_n, \dot{x}_2) & \cdots & \text{Cov}(\dot{x}_n, \dot{x}_{n-1}) & \text{Var}(\dot{x}_n) \end{pmatrix}\end{aligned}$$

Con esto en mente, es fácil ver que:

$$\begin{aligned}\text{Var}\left(f\left(\overset{\circ}{\vec{x}}\right)\right) &= \sum_{i=1}^n \sum_{j=1}^n \left(\vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right)_i \left(\vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right)_j \left(\text{Cov}\left(\overset{\circ}{\vec{x}}\right)\right)_{i,j} + o(3) = \\ &= \sum_{i=1}^n \left(\vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right)_i \left(\sum_{j=1}^n \left(\text{Cov}\left(\overset{\circ}{\vec{x}}\right)\right)_{i,j} \left(\vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right)_j\right) + o(3) = \\ &= \sum_{i=1}^n \left(\vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right)_i \left(\text{Cov}\left(\overset{\circ}{\vec{x}}\right) \vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right)_i = \left[\vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right)\right]^T \text{Cov}\left(\overset{\circ}{\vec{x}}\right) \vec{\nabla}f\left(\text{E}\left(\overset{\circ}{\vec{x}}\right)\right) + o(3)\end{aligned}$$

donde hemos usado los subíndices i y j para indicar los elementos de las matrices o vectores correspondientes y T indica «transpuesta».

Q.E.D.

Corolario 2.1. Sea Ω un abierto de \mathbb{R}^n . A continuación, sea $\vec{x} = (\dot{x}_1, \dots, \dot{x}_n)$ un vector de variables aleatorias independientes dos a dos y sea $f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una función escalar de clase $C^{(1)}$. Al primer orden no nulo, se tiene:

$$\mathbb{E} \left[f \left(\vec{x} \right) \right] = f \left(\mathbb{E} \left(\vec{x} \right) \right) + o(2)$$

$$\text{Var} \left[f \left(\vec{x} \right) \right] = \sum_{i=1}^n \left[\frac{\partial f}{\partial \dot{x}_i} \left(\mathbb{E} \left(\vec{x} \right) \right) \right]^2 \text{Var} (\dot{x}_i) + o(3)$$

Demostración. El primer resultado se obtiene directamente del teorema 2.1 en la página 10. Por ese mismo teorema se tiene:

$$\text{Var} \left(f \left(\vec{x} \right) \right) = \left[\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \right]^T \text{Cov} \left(\vec{x} \right) \vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) + o(3)$$

En nuestro caso, como las variables $\vec{x} = (\dot{x}_1, \dots, \dot{x}_n)$ son independientes dos a dos, se tiene:

$$\text{Cov} (\dot{x}_i, \dot{x}_j) = 0 \quad \forall i \neq j \wedge i, j = 1, \dots, n$$

Por tanto, la matriz $\text{Cov} \left(\vec{x} \right)$ es diagonal. De esta forma, el producto vector*matriz*vector queda, sencillamente:

$$\begin{aligned} \text{Var} \left(f \left(\vec{x} \right) \right) &= \sum_{i=1}^n \sum_{j=1}^n \left(\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \right)_i \left(\text{Cov} \left(\vec{x} \right) \right)_{i,j} \left(\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \right)_j = \\ &= \sum_{i=1}^n \left[\left(\vec{\nabla} f \left(\mathbb{E} \left(\vec{x} \right) \right) \right)_i \right]^2 \left(\text{Cov} \left(\vec{x} \right) \right)_{i,i} \end{aligned}$$

ya que $\text{Cov} (\dot{x}_i, \dot{x}_j) = 0 \forall i \neq j$. De esta forma, recordando que $\left(\text{Cov} \left(\vec{x} \right) \right)_{i,i} = \text{Var} (\dot{x}_i)$ y que $\left(\vec{\nabla} f \right)_i = \frac{\partial f}{\partial \dot{x}_i}$, se llega a:

$$\text{Var} \left(f \left(\vec{x} \right) \right) = \sum_{i=1}^n \left[\frac{\partial f}{\partial \dot{x}_i} \left(\mathbb{E} \left(\vec{x} \right) \right) \right]^2 \text{Var} (\dot{x}_i)$$

Q.E.D.

3. Ajustes lineales con errores únicamente en las ordenadas

Suposición 3.1. Los datos medidos $\{x_i\}_{i=1}^n$ son exactos.

Suposición 3.2. Cada dato medido y_i es una muestra de una distribución gaussiana de centro «el valor real» \tilde{y}_i desconocido y de varianza σ_i^2 conocida, que denotaremos \hat{y}_i . Es decir:

$$\hat{y}_i \sim \mathcal{N} (\tilde{y}_i, \sigma_i^2) \quad \forall i = 1, \dots, n$$

Además, las variables aleatorias asociadas a los datos medidos $\{\hat{y}_i\}_{i=1}^n$ son independientes dos a dos.

Suposición 3.3. Los valores reales $\{\tilde{y}_i\}_{i=1}^n$ satisfacen la ecuación:

$$\tilde{y}_i = ax_i + b \quad \forall i = 1, \dots, n$$

donde $a, b \in \mathbb{R}$ son dos constantes (los parámetros del modelo).

Proposición 3.1. Los estimadores de máxima verosimilitud para los parámetros a y b del modelo vienen dados por:

$$\hat{a} = \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2}\right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right)}$$

$$\hat{b} = \frac{\left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) \hat{a}}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)}$$

Demostración. Por la suposición 3.2 en la página anterior, sabemos que los datos medidos $\{y_i\}_{i=1}^n$ son muestras de variables aleatorias gaussianas de centro \tilde{y}_i y desviación típica σ_i . Es decir:

$$\dot{y}_i \sim \mathcal{N}(\tilde{y}_i, \sigma_i^2) \quad \forall i = 1, \dots, n$$

Ahora bien, por la suposición 3.3 en la página anterior, se tiene $\tilde{y}_i = ax_i + b$. Entonces, las variables aleatorias $\{\dot{y}_i\}_{i=1}^n$ siguen distribuciones normales de centro $ax_i + b$ y de desviación típica σ_i . Dicho de otra forma:

$$\dot{y}_i \sim \mathcal{N}(ax_i + b, \sigma_i^2) \quad \forall i = 1, \dots, n$$

Sabemos que lo anterior está bien definido ya que los $\{x_i\}_{i=1}^n$ son exactos por la suposición 3.1 en la página anterior.

De esta manera, la función de densidad de probabilidad de cada variable aleatoria \dot{y}_i viene dada por:

$$f_{\dot{y}_i}(\dot{y}_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(\dot{y}_i - ax_i - b)^2}{2\sigma_i^2}\right) \quad \forall i = 1, \dots, n$$

donde no debe confundirse la variable aleatoria \dot{y}_i con el argumento de la función de densidad de probabilidad $f_{\dot{y}_i}: \dot{y}_i$.

A continuación como, por la suposición 3.2 en la página anterior, las variables aleatorias $\{\dot{y}_i\}_{i=1}^n$ son independientes dos a dos, la función de verosimilitud será simplemente el producto de las funciones de densidad $f_{\dot{y}_i}$ evaluadas en los datos medidos:

$$L(a, b) = \prod_{i=1}^n f_{\dot{y}_i}(\dot{y}_i = y_i) = \frac{1}{\sqrt{2^n \pi^n \prod_{i=1}^n \sigma_i^2}} \exp\left(-\sum_{i=1}^n \left(\frac{y_i - ax_i - b}{\sqrt{2}\sigma_i}\right)^2\right)$$

donde hemos usado la propiedad $e^a e^b = e^{a+b} \forall a, b \in \mathbb{R}$. Para proseguir, consideramos el logaritmo de la función de verosimilitud:

$$l(a, b) = \ln L = \ln\left(\frac{1}{\sqrt{2^n \pi^n \prod_{i=1}^n \sigma_i^2}}\right) - \sum_{i=1}^n \left(\frac{y_i - ax_i - b}{\sqrt{2}\sigma_i}\right)^2$$

donde hemos hecho uso de la propiedad del logaritmo de transformar productos en sumas y hemos empleado, asimismo, que la exponencial y el logaritmo son operaciones inversas.

Con el fin de obtener los operadores de máxima verosimilitud, debemos buscar aquellos puntos en los que l es máxima (que ya sabemos que coincidirán con los de L por ser el logaritmo neperiano una función estrictamente creciente). Para ello, derivamos la función l con respecto a los parámetros a y b :

$$\frac{\partial l}{\partial a} = - \sum_{i=1}^n 2 \frac{y_i - ax_i - b}{\sqrt{2}\sigma_i} \left(-\frac{x_i}{\sqrt{2}\sigma_i} \right) = \sum_{i=1}^n x_i \frac{y_i - ax_i - b}{\sigma_i^2} = \sum_{i=1}^n x_i \frac{y_i - ax_i - b}{\sigma_i^2} \quad (3.1)$$

$$\frac{\partial l}{\partial b} = - \sum_{i=1}^n 2 \frac{y_i - ax_i - b}{\sqrt{2}\sigma_i} \left(-\frac{1}{\sqrt{2}\sigma_i} \right) = \sum_{i=1}^n \frac{y_i - ax_i - b}{\sigma_i^2} = \sum_{i=1}^n \frac{y_i - ax_i - b}{\sigma_i^2} \quad (3.2)$$

donde hemos hecho uso de que la derivada parcial es una aplicación lineal y de la regla de la cadena. Ahora, imponemos $\vec{\nabla}l = \left(\frac{\partial l}{\partial a}, \frac{\partial l}{\partial b} \right) = \vec{0}$ para obtener el punto crítico de l :

$$\frac{\partial l}{\partial b} = 0 \iff \sum_{i=1}^n x_i \frac{y_i - ax_i - b}{\sigma_i^2} = 0$$

$$\frac{\partial l}{\partial a} = 0 \iff \sum_{i=1}^n \frac{y_i - ax_i - b}{\sigma_i^2} = 0$$

Expandiendo las sumas anteriores y sacando los parámetros a y b como factor común de los sumatorios, se tiene:

$$\frac{\partial l}{\partial b} = 0 \iff \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) a - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) b = 0 \quad (3.3)$$

$$\frac{\partial l}{\partial a} = 0 \iff \left(\sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right) a - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) b = 0 \quad (3.4)$$

Nuestro objetivo es obtener los valores de a y b que resuelven el sistema de ecuaciones anterior. Para tal fin, aplicaremos el método de la reducción, buscando que el coeficiente de b sea el mismo en ambas ecuaciones. Por consiguiente, multiplicamos la ecuación 3.3 por $\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)$ y la ecuación 3.4 por $\left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right)$. De esta forma, se obtiene:

$$\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 a - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) b = 0 \quad (3.5)$$

$$\left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right) a - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) b = 0 \quad (3.6)$$

Nótese que, como estábamos buscando, ahora el coeficiente que acompaña a b en ambas ecuaciones es el mismo. De esta forma, restándole a la ecuación 3.5 la ecuación 3.6, se tiene:

$$\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} \right) - \left[\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right) \right] a = 0$$

ya que el término en b desaparece. De la ecuación anterior resulta razonablemente sencillo despejar a :

$$\begin{aligned} \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} \right) &= \left[\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right) \right] a \iff \\ \iff a &= \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} \right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right)} \end{aligned}$$

Conocido el valor de a , podemos simplemente despejar b de la ecuación 3.3 en la página anterior:

$$\begin{aligned} \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) a - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) b &= 0 \iff \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) a = \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) b \iff \\ &\iff b = \frac{\left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) a}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right)} \end{aligned}$$

Por tanto, los estimadores de máxima verosimilitud para los operadores a y b serán:

$$\begin{aligned} \hat{a} &= \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2} \right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right)} \\ \hat{b} &= \frac{\left(\sum_{i=1}^n \frac{y_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \hat{a}}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right)} \end{aligned}$$

Q.E.D.

Proposición 3.2. *Los estimadores \hat{a} y \hat{b} son no sesgados.*

Demostración. Recordemos que un operador \hat{a} es no sesgado si $E(\hat{a}) = a$. Para este cálculo, consideraremos que las muestras y_i que aparecen en las expresiones de la proposición 3.1 en la página 13 representan sus variables aleatorias asociadas \mathring{y}_i .

$$\hat{a} = \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{\mathring{y}_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i \mathring{y}_i}{\sigma_i^2} \right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right)}$$

Como, por la suposición 3.1 en la página 12, los $\{x_i\}_{i=1}^n$ son datos exactos, se tiene:

$$\begin{aligned} E(\hat{a}) &= \frac{E \left[\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{\mathring{y}_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i \mathring{y}_i}{\sigma_i^2} \right) \right]}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right)} = \\ &= \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) E \left(\sum_{i=1}^n \frac{\mathring{y}_i}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) E \left(\sum_{i=1}^n \frac{x_i \mathring{y}_i}{\sigma_i^2} \right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right)} = \\ &= \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{E(\mathring{y}_i)}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i E(\mathring{y}_i)}{\sigma_i^2} \right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right)} \end{aligned}$$

A continuación, combinando las suposiciones 3.2 en la página 12 y 3.3 en la página 12, se tiene que:

$$E(\mathring{y}_i) = \tilde{y}_i = ax_i + b \quad \forall i = 1, \dots, n \quad (3.7)$$

donde los parámetros a y b que aparecen en la ecuación anterior no son sus estimadores, sino sus valores verdaderos (por eso no llevan circunflejo). De esta forma, se tiene:

$$E(\hat{a}) = \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{ax_i+b}{\sigma_i^2} \right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i(ax_i+b)}{\sigma_i^2} \right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2} \right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \right)}$$

Operando, llegamos a:

$$\begin{aligned}
E(\hat{a}) &= \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) \left(a \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) + b \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)\right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(a \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right) + b \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)\right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right)} = \\
&= \frac{a \left[\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right)\right] + b \left[\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)\right]}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right)} = \\
&= a \frac{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right)}{\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)^2 - \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right)} = a
\end{aligned}$$

De esta forma, efectivamente $E(\hat{a}) = a$ y, por tanto, \hat{a} es un operador no sesgado. Vayamos con \hat{b} :

$$\hat{b} = \frac{\left(\sum_{i=1}^n \frac{\hat{y}_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) \hat{a}}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)}$$

De nuevo, para calcular el valor esperado de \hat{b} , tenemos que considerar que todo es constante, salvo las variables aleatorias $\{\hat{y}_i\}_{i=1}^n$ y el estimador \hat{a} (que también es una variable aleatoria). De esta manera:

$$\begin{aligned}
E(\hat{b}) &= \frac{E\left[\left(\sum_{i=1}^n \frac{\hat{y}_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) \hat{a}\right]}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)} = \frac{E\left(\sum_{i=1}^n \frac{\hat{y}_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) E(\hat{a})}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)} = \\
&= \frac{\left(\sum_{i=1}^n \frac{E(\hat{y}_i)}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) E(\hat{a})}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)}
\end{aligned}$$

Al igual que en el caso anterior, combinando las suposiciones 3.2 en la página 12 y 3.3 en la página 12, se tiene la ecuación 3.7 en la página anterior. Además, como hemos probado antes $E(\hat{a}) = a$. Así, se obtiene:

$$\begin{aligned}
E(\hat{b}) &= \frac{\left(\sum_{i=1}^n \frac{ax_i+b}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) a}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)} = \\
&= \frac{a \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) + b \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) a}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)} = \\
&= \frac{a \left[\left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)\right] + b \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)} = b \frac{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)}{\left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right)} = b
\end{aligned}$$

Y, por consiguiente, $E(\hat{b}) = b$ y \hat{b} es un operador no sesgado.

Q.E.D.

Proposición 3.3. La matriz de covarianza de los estimadores \hat{a}, \hat{b} satisface:

$$\text{Cov}(\hat{a}, \hat{b}) \geq \frac{1}{\left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)^2} \begin{pmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} & -\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \\ -\sum_{i=1}^n \frac{x_i}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \end{pmatrix}$$

donde la desigualdad \geq se entiende como $A \geq B \iff A - B$ es semidefinida positiva.

Demostración. En la demostración de la proposición 3.1 en la página 13 llegamos a las siguientes expresiones para las derivadas parciales del logaritmo de la función de verosimilitud (ecuaciones 3.1 en la página 14 y 3.2 en la página 14):

$$\frac{\partial l}{\partial a} = \sum_{i=1}^n x_i \frac{y_i - ax_i - b}{\sigma_i^2}$$

$$\frac{\partial l}{\partial b} = \sum_{i=1}^n \frac{y_i - ax_i - b}{\sigma_i^2}$$

A continuación, debemos hallar las derivadas segundas:

$$\frac{\partial^2 l}{\partial a^2} = -\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \quad \frac{\partial^2 l}{\partial b^2} = -\sum_{i=1}^n \frac{1}{\sigma_i^2} \quad \frac{\partial^2 l}{\partial a \partial b} = -\sum_{i=1}^n \frac{x_i}{\sigma_i^2}$$

De esta forma, el hessiano del logaritmo de la función de verosimilitud queda:

$$Hl = \begin{pmatrix} -\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} & -\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \\ -\sum_{i=1}^n \frac{x_i}{\sigma_i^2} & -\sum_{i=1}^n \frac{1}{\sigma_i^2} \end{pmatrix}$$

Y, por ende, la matriz de información de Fisher resulta:

$$I = E[-Hl] = \begin{pmatrix} \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i}{\sigma_i^2} \\ \sum_{i=1}^n \frac{x_i}{\sigma_i^2} & \sum_{i=1}^n \frac{1}{\sigma_i^2} \end{pmatrix}$$

Sabemos que la matriz de información de Fisher satisface la desigualdad:

$$\text{Cov}(\hat{a}, \hat{b}) \geq I^{-1} = \frac{1}{\left(\sum_{i=1}^n \frac{x_i^2}{\sigma_i^2}\right) \left(\sum_{i=1}^n \frac{1}{\sigma_i^2}\right) - \left(\sum_{i=1}^n \frac{x_i}{\sigma_i^2}\right)^2} \begin{pmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} & -\sum_{i=1}^n \frac{x_i}{\sigma_i^2} \\ -\sum_{i=1}^n \frac{x_i}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \end{pmatrix}$$

donde hemos aplicado que la inversa de una matriz 2x2 viene dada por la fórmula:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Q.E.D.

Proposición 3.4. *La variable aleatoria:*

$$\hat{Z} = \sum_{i=1}^n \frac{(\hat{y}_i - ax_i - b)^2}{\sigma_i^2}$$

sigue una distribución χ^2 con $n - 2$ grados de libertad.

Demostración. Por la suposición 3.2 en la página 12, $\hat{y}_i \sim \mathcal{N}(\tilde{y}_i, \sigma_i) \quad \forall i = 1, \dots, n$. Además, por la suposición 3.3 en la página 12, se tiene:

$$\hat{y}_i \sim \mathcal{N}(ax_i + b, \sigma_i) \quad \forall i = 1, \dots, n$$

Por tanto:

$$\hat{\varepsilon}_i := \frac{\hat{y}_i - (ax_i + b)}{\sigma_i} = \frac{\hat{y}_i - ax_i - b}{\sigma_i} \sim \mathcal{N}(0, 1) \quad \forall i = 1, \dots, n$$

por la proposición 1.3 en la página 4. Por otra parte, las expresiones de \hat{a} y \hat{b} obtenidas en la proposición 3.1 en la página 13 imponen dos ligaduras en los posibles valores de las variables aleatorias $\{\hat{\varepsilon}_i\}_{i=1}^n$. Por tanto, el vector aleatorio $\hat{\varepsilon} := (\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n)$ pertenece a un espacio vectorial de dimensión $n - 2$. Por la proposición 1.5 en la página 9:

$$\hat{Z} = \sum_{i=1}^n \frac{(\hat{y}_i - ax_i - b)^2}{\sigma_i^2} = \sum_{i=1}^n \hat{\varepsilon}_i^2 = \|\hat{\varepsilon}\|^2 \sim \chi_{n-2}^2$$

Q.E.D.

Implementación en el ordenador 3.1. Para implementar todo esto en el ordenador resulta más sencillo definirse las cantidades:

$$S_\sigma := \sum_{i=1}^n \frac{1}{\sigma_i^2} \quad S_x := \sum_{i=1}^n \frac{x_i}{\sigma_i^2} \quad S_{xx} := \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \quad S_y := \sum_{i=1}^n \frac{y_i}{\sigma_i^2} \quad S_{xy} := \sum_{i=1}^n \frac{x_i y_i}{\sigma_i^2}$$

De esta forma, se obtienen las siguientes expresiones:

$$\hat{a} = \frac{S_x S_y - S_\sigma S_{xy}}{S_x^2 - S_\sigma S_{xx}}$$

$$\hat{b} = \frac{S_y - S_x \hat{a}}{S_\sigma}$$

$$\text{Cov}(\hat{a}, \hat{b}) \geq \frac{1}{S_{xx} S_\sigma - S_x^2} \begin{pmatrix} S_\sigma & -S_x \\ -S_x & S_{xx} \end{pmatrix}$$

En la práctica, la desigualdad anterior se interpreta como una igualdad por comodidad.

4. Ajustes a funciones arbitrarias con errores únicamente en las ordenadas

Suposición 4.1. Los datos medidos $\{x_i\}_{i=1}^n$ son exactos.

Suposición 4.2. Cada dato medido y_i es una muestra de una distribución gaussiana de centro «el valor real» \tilde{y}_i desconocido y de varianza σ_i^2 conocida, que denotaremos \hat{y}_i . Es decir:

$$\hat{y}_i \sim \mathcal{N}(\tilde{y}_i, \sigma_i^2) \quad \forall i = 1, \dots, n$$

Además, las variables aleatorias asociadas a los datos medidos $\{\hat{y}_i\}_{i=1}^n$ son independientes dos a dos.

Suposición 4.3. Los valores reales $\{\tilde{y}_i\}_{i=1}^n$ satisfacen la ecuación:

$$\tilde{y}_i = g(x_i, a_1, \dots, a_m) \quad \forall i = 1, \dots, n$$

donde $a_1, \dots, a_m \in \mathbb{R}$ son constantes (los parámetros del modelo) y g es una función $g : \Omega \subseteq \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$, siendo Ω un conjunto abierto de \mathbb{R}^{m+1} .

Proposición 4.1. Los estimadores de máxima verosimilitud para los parámetros del modelo $(\hat{a}_1, \dots, \hat{a}_m)$ son los valores de a_1, \dots, a_m que resuelven el sistema de ecuaciones:

$$\frac{\partial l}{\partial a_j} = \sum_{i=1}^n \frac{y_i - g(x_i, a_1, \dots, a_m)}{\sigma_i^2} \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) = 0 \quad \forall j = 1, \dots, m$$

Demostración. Por la suposición 4.2, sabemos que los datos medidos $\{y_i\}_{i=1}^n$ son muestras de variables aleatorias gaussianas de centro \tilde{y}_i y desviación típica σ_i . Es decir:

$$\hat{y}_i \sim \mathcal{N}(\tilde{y}_i, \sigma_i^2) \quad \forall i = 1, \dots, n$$

Ahora bien, por la suposición 4.3 en la página anterior, se tiene $\tilde{y}_i = g(x_i, a_1, \dots, a_m)$. Entonces, las variables aleatorias $\{\tilde{y}_i\}_{i=1}^n$ siguen distribuciones normales de centro $g(x_i, a_1, \dots, a_m)$ y de desviación típica σ_i . Dicho de otra forma:

$$\tilde{y}_i \sim \mathcal{N}(g(x_i, a_1, \dots, a_m), \sigma_i^2) \quad \forall i = 1, \dots, n$$

Sabemos que lo anterior está bien definido ya que los $\{x_i\}_{i=1}^n$ son exactos por la suposición 4.1 en la página anterior.

De esta manera, la función de densidad de probabilidad de cada variable aleatoria \tilde{y}_i viene dada por:

$$f_{\tilde{y}_i}(\tilde{y}_i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(\tilde{y}_i - g(x_i, a_1, \dots, a_m))^2}{2\sigma_i^2}\right) \quad \forall i = 1, \dots, n$$

donde no debe confundirse la variable aleatoria \tilde{y}_i con el argumento de la función de densidad de probabilidad $f_{\tilde{y}_i}: \tilde{y}_i$.

A continuación, como por la suposición 4.2 en la página anterior, las variables aleatorias $\{\tilde{y}_i\}_{i=1}^n$ son independientes dos a dos, la función de verosimilitud será simplemente el producto de las funciones de densidad $f_{\tilde{y}_i}$ evaluadas en los datos medidos:

$$L(a, b) = \prod_{i=1}^n f_{\tilde{y}_i}(\tilde{y}_i = y_i) = \frac{1}{\sqrt{2^n \pi^n \prod_{i=1}^n \sigma_i^2}} \exp\left(-\sum_{i=1}^n \left(\frac{y_i - g(x_i, a_1, \dots, a_m)}{\sqrt{2}\sigma_i}\right)^2\right)$$

donde hemos usado la propiedad $e^{ae^b} = e^{a+b} \forall a, b \in \mathbb{R}$. Para proseguir, consideramos el logaritmo de la función de verosimilitud:

$$l(a, b) = \ln L = \ln\left(\frac{1}{\sqrt{2^n \pi^n \prod_{i=1}^n \sigma_i^2}}\right) - \sum_{i=1}^n \left(\frac{y_i - g(x_i, a_1, \dots, a_m)}{\sqrt{2}\sigma_i}\right)^2$$

donde hemos hecho uso de la propiedad del logaritmo de transformar productos en sumas y hemos empleado, asimismo, que la exponencial y el logaritmo son operaciones inversas.

Con el fin de obtener los operadores de máxima verosimilitud, debemos buscar aquellos puntos en los que l es máxima (que ya sabemos que coincidirán con los de L por ser el logaritmo neperiano una función estrictamente creciente). Para ello, derivamos la función l con respecto a cada parámetro a_j :

$$\begin{aligned} \frac{\partial l}{\partial a_j} &= -\sum_{i=1}^n \frac{y_i - g(x_i, a_1, \dots, a_m)}{\sqrt{2}\sigma_i} \left(-\frac{1}{\sqrt{2}\sigma_i}\right) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) = \\ &= \sum_{i=1}^n \frac{y_i - g(x_i, a_1, \dots, a_m)}{\sigma_i^2} \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) \end{aligned}$$

Los estimadores de máxima verosimilitud $\hat{a}_1, \dots, \hat{a}_m$ son los valores de a_1, \dots, a_m tales que:

$$\frac{\partial l}{\partial a_j} = 0 \quad \forall j = 1, \dots, m \iff$$

$$\sum_{i=1}^n \frac{y_i - g(x_i, a_1, \dots, a_m)}{\sigma_i^2} \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) = 0 \quad \forall j = 1, \dots, m$$

Q.E.D.

Proposición 4.2. *La matriz de información de Fisher viene dada por:*

$$I_{j,k} = \sum_{i=1}^n \frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k}(x_i, a_1, \dots, a_m) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) \quad \forall j, k = 1, \dots, m$$

Demostración. Recuperamos la expresión para la derivada del logaritmo de la función de verosimilitud respecto a un parámetro a_j dada en la proposición 4.1 en la página 18:

$$\frac{\partial l}{\partial a_j} = \sum_{i=1}^n \frac{y_i - g(x_i, a_1, \dots, a_m)}{\sigma_i^2} \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m)$$

Volviendo a derivar con respecto a a_k , se tiene:

$$\begin{aligned} \frac{\partial^2 l}{\partial a_k \partial a_j} = & \sum_{i=1}^n \left[-\frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k}(x_i, a_1, \dots, a_m) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) + \right. \\ & \left. + \frac{y_i - g(x_i, a_1, \dots, a_m)}{\sigma_i^2} \frac{\partial^2 g}{\partial a_k \partial a_j}(x_i, a_1, \dots, a_m) \right] \end{aligned}$$

Ahora bien, sabemos que el elemento j, k de la matriz de información de Fisher viene dado por:

$$I_{j,k} = E \left(-(\text{H}l)_{j,k} \right) = E \left(-\frac{\partial^2 l}{\partial a_k \partial a_j} \right)$$

donde en la parcial segunda correspondiente hay que sustituir y_i por \hat{y}_i , su variable aleatoria correspondiente.

$$\begin{aligned} I_{j,k} = E & \left[\sum_{i=1}^n \left[\frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k}(x_i, a_1, \dots, a_m) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) + \right. \right. \\ & \left. \left. - \frac{\hat{y}_i - g(x_i, a_1, \dots, a_m)}{\sigma_i^2} \frac{\partial^2 g}{\partial a_k \partial a_j}(x_i, a_1, \dots, a_m) \right] \right] = \\ & = \sum_{i=1}^n \frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k}(x_i, a_1, \dots, a_m) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) + \\ & - \sum_{i=1}^n \frac{E(\hat{y}_i) - g(x_i, a_1, \dots, a_m)}{\sigma_i^2} \frac{\partial^2 g}{\partial a_k \partial a_j}(x_i, a_1, \dots, a_m) \end{aligned}$$

Combinando las suposiciones 4.2 en la página 18 y 4.3 en la página 18, se tiene que:

$$E(\hat{y}_i) = g(x_i, a_1, \dots, a_m) \quad \forall i = 1, \dots, n$$

Y, por tanto, el elemento j, k de la matriz de información de Fisher queda, simplemente:

$$I_{j,k} = \sum_{i=1}^n \frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k}(x_i, a_1, \dots, a_m) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m)$$

Q.E.D.

Proposición 4.3. *La variable aleatoria:*

$$\hat{Z} = \sum_{i=1}^n \frac{(\hat{y}_i - g(x_i, a_1, \dots, a_m))^2}{\sigma_i^2}$$

sigue una distribución χ^2 con $n - m$ grados de libertad.

Demostración. Por la suposición 4.2 en la página 18, $\hat{y}_i \sim \mathcal{N}(\tilde{y}_i, \sigma_i^2) \quad \forall i = 1, \dots, n$. Además, por la suposición 4.3 en la página 18, se tiene:

$$\hat{y}_i \sim \mathcal{N}(g(x_i, a_1, \dots, a_m), \sigma_i^2) \quad \forall i = 1, \dots, n$$

Por tanto:

$$\hat{\varepsilon}_i := \frac{\hat{y}_i - g(x_i, a_1, \dots, a_m)}{\sigma_i} \sim \mathcal{N}(0, 1) \quad \forall i = 1, \dots, n$$

por la proposición 1.3 en la página 4. Por otra parte, las ecuaciones obtenidas en la proposición 4.1 en la página 18 imponen m ligaduras en los posibles valores de las variables aleatorias $\{\hat{\varepsilon}_i\}_{i=1}^n$. Por tanto, el vector aleatorio $\hat{\varepsilon} := (\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n)$ pertenece a un espacio vectorial de dimensión $n - m$. Por la proposición 1.5 en la página 9:

$$\hat{Z} = \sum_{i=1}^n \frac{(\hat{y}_i - g(x_i, a_1, \dots, a_m))^2}{\sigma_i^2} = \sum_{i=1}^n \hat{\varepsilon}_i^2 = \|\hat{\varepsilon}\|^2 \sim \chi_{n-m}^2$$

Q.E.D.

Implementación en el ordenador 4.1. *Este caso resulta más complicado que los ajustes lineales puesto que no hay solución analítica. El método a seguir es el siguiente:*

1. Se calculan analíticamente (si se puede) las derivadas de la función g : $\frac{\partial g}{\partial x_j} \quad \forall j = 1, \dots, m$.
2. Se resuelven numéricamente las ecuaciones 4.1 en la página 18.

$$\frac{\partial l}{\partial a_j} = \sum_{i=1}^n \frac{y_i - g(x_i, a_1, \dots, a_m)}{\sigma_i^2} \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) = 0 \quad \forall j = 1, \dots, m$$

3. Se calcula la matriz de información de Fisher mediante la expresión dada en la proposición 4.2 en la página anterior:

$$I_{j,k} = \sum_{i=1}^n \frac{1}{\sigma_i^2} \frac{\partial g}{\partial a_k}(x_i, a_1, \dots, a_m) \frac{\partial g}{\partial a_j}(x_i, a_1, \dots, a_m) \quad \forall j, k = 1, \dots, m$$

4. Se invierte la matriz de información de Fisher numéricamente para obtener la matriz de covarianza:

$$\text{Cov}(\hat{a}_1, \dots, \hat{a}_m) = I^{-1}$$