Considerando un conjunto $y_1, y_2, ..., y_n$ de variables aleatorias independientes relacionadas con otras variables x_i que se asume conocida sin error.

Cada y_i tiene un valor medio λ_i desconocido y una varianza σ_i^2 conocida. Las N medidas de y_i pueden considerarse como las medidas de un vector aleatorio N-dimensional con función de distribución de probabilidad

$$g(y_1, ..., y_n; \lambda_1, ..., \lambda_n, \sigma_1, ..., \sigma_n) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(y_i - \lambda_i)^2}{2\sigma_i^2}\right].$$

Sse debe de suponer además que el valor verdadero de las y_i es una función de la variable x que depende de un vector de parámetros desconocidos en principio.

$$\lambda = \lambda(x_{i;\theta}) \qquad\qquad \theta = (\theta_1, ..., \theta_m).$$

El objetivo de mínimos cuadrados es estimar el vector de parámetros θ. Para establecer el método tomamos logaritmos en la función de distribución de probabilidad que describe los datos.

$$\begin{aligned} Ln(g) &= A + Ln(\theta) \qquad\qquad donde \\ A &= \sum_{i=1}^{N} Ln\left[\frac{1}{\sqrt{2\pi\sigma_i^2}}\right] \\ Ln[L(\theta)] &= -\frac{1}{2}\sum_{i=1}^{N} \frac{(y_i - \lambda_i)^2}{2\sigma_i^2} \end{aligned}$$

Para obtener el mejor ajuste se debe minimizar la cantidad χ^2 definida por:

$$\chi^2(\theta) = \sum_{i=1}^{N} \frac{(y_i - \lambda_i)^2}{2\sigma_i^2}$$

Para minimizarla se debe cumplir que $\dfrac{d\chi^2}{d\theta} = 0$, entonces

$$\frac{d\chi^2}{d\theta} = \left[\frac{d\lambda}{d\theta}\right]^2 - 2y_i\frac{d\lambda}{d\theta} = 0,$$

resolviendo la ecuación cuadrática se tiene que:

$$\frac{d\lambda}{d\theta} = y_i \pm y_i,$$

el mínimo se obtiene para $2y_i$, por lo que

$$\lambda(x_i, \theta) = ax_i + b.$$

Donde se a, b son constantes. Se ha tomado el hecho de que las y_i están relacionadas con las x_i. Así se ha mostrado que el método de mínimos cuadrados para una recta se obtiene de la minimización de la χ^2.