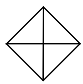

Why

We consider the probabilistic linear model in which all random variables are normal.

Definition

A *normal linear model* is a probabilistic linear model in which the parameter and noise vectors have normal (Gaussian) densities. The model is also called the *Gaussian linear model* or the *linear model with Gaussian noise*.

Let $(x : \Omega \to \mathbf{R}^d, A \in \mathbf{R}^{n \times d}, e : \Omega \to \mathbf{R}^n)$ be a probabilistic linear model over the probability space $(\Omega, \mathcal{A}, \mathbf{P})$ in which x and e have normal densities. Recall that a probabilistic linear model has observation vector $y : \Omega \to \mathbf{R}^n$ defined by

$$y = Ax + e.$$

Conditional density of x on y

Since x and e are normal and independent, y is normal.[1] Moreover, the random vector (x, y) is normal with covariance

$$\begin{pmatrix} \Sigma_x & \Sigma_x A^\top \\ A\Sigma_x & A\Sigma_x A^\top + \Sigma_e \end{pmatrix}.$$

So the conditional density (see Normal Conditionals) of $g_{x|y}(\cdot, \gamma)$ is normal with mean

$$\Sigma_x A^\top \left(A\Sigma_x A^\top + \Sigma_e\right)^{-1} \gamma$$

and covariance

$$\Sigma_x - \Sigma_x A^\top \left(A\Sigma_x A^\top + \Sigma_e\right)^{-1} A\Sigma_x.$$

This density is sometimes called the posterior for the parameters given the observations. So the parameter posterior of the normal linear model is normal.

[1] Future editions will include an account.

We can write the conditional mean as

$$\left(\Sigma_x{}^{-1} + A^\top \Sigma_e{}^{-1} A\right)^{-1} A^\top \Sigma_e{}^{-1}$$

and the conditional covariance as

$$\left(\Sigma_x{}^{-1} + A^\top \Sigma_e{}^{-1} A\right)^{-1}.$$ [2]

Very frequently we use these forms when $d < n$. In other words, in the case that we have fewer unknowns than measurements. In that case Σ_x is smaller than $A\Sigma_x A^\top$.

Maximum conditional estimate of x

Proposition 1. *The maximum conditional estimate of* $x : \Omega \to \mathbf{R}^d$ *given observed value* $\gamma \in \mathbf{R}^n$ *of* $y : \Omega \to \mathbf{R}^n$ *is the conditional mean*

$$\Sigma_x A^\top \left(A\Sigma_x A^\top + \Sigma_e\right)^{-1} \gamma.$$

[2] A proof will appear in future editions. Use the matrix inversion lemma or facts about inverses.