Gauss markov assumptions time series

It is not confused with the Gauss-Markov process. Blue redirects here. For the queue management algorithm, look at Blue (queue management algorithm). Part of a series onRegression analysis of linear regression models simple multimodal regression multifaceted linear model general linear model discrete linear model selection bipolar regression bipolar regression logistic regression Logit mixed logit prebit multi-procedural ordered Logit order Probit Poisson multilevel model fixed random effects mixed linear effects - effects of nonlinear mixed model effects nonlinear semipapermometric semipapermometric semi-parametric semi-parametric model Segmented in the variable of estimating the minimum nonlinear linear square of ordinary weighted ordinary partial total non-negative regression ridge regular minimum absolute deviations repeatedly weight of bayesian multivariate background validation mean and prediction of response errors and remaining good of student fit remaining Gauss remnants - Markov portalvte mathematics theory in statistics, the Gauss–Markov theorem (or simply Gauss theorem for some authors) [1] states that the ordinary least squares (OLS) estimator has The lowest sampling variance within the class of linear unbiased estimators, if errors in the linear regression model are not linked, have equal variance and zero waiting value. [2] Errors do not require normality, nor do they require the same independentness and distribution (they are not associated with the average zero and hemostic with limited variance only). The condition that the estimator is impartial cannot be reduced, as there are biased estimates with lower variances. For example, see the James-Stein estimator (which also abandons linearity), the mane regression, or simply any staple estimator. The theory was named after Karl Friedrich Gauss and Andrei Markov, although Gauss's work is significantly ahead of Markov. [3] However, while Gauss achieved the result assuming independence and normality, Markov reduced the assumptions to the form expressed above. [4] A further generalization of non-spherical errors was given by Alexander Aiken. [5] The statement assumes we are referring to matrix, y_ = X β _ + ε _ , (y _ , ε _ ∈ R n , β _ ∈ R K and X ∈ R n × K) {\displaystyle {\underline {y}}=X{\underline {\beta}+{underline {\varepsilon},\quad ({\underline {y}},{\underline {\varepsilon}}}\in \mathbb {R} ^{n},{\underline {\beta}}\in\mathbb {R} ^{K}{text{ and }}X\in \mathbb {R} ^{n\times K})} expanding to , y i = ∑ j = 1 K β j x i j + ε i ∀ i = 1 , 2 , ... , n {\displaystyle y_{i}=\sum_{j=1}^{K}\beta_{{{ j}X_{ij}+\varepsilon_{i}\quad\forall i=1,2,\ldots,n} where β j {\displaystyle \beta_{j}} are non-random but unobservable parameters, X j {\displaystyle X_{ij}} are non-random and visible (called variables), ε i {\displaystyle \varepsilon _{i}} are random, and so y i {\displaystyle y_{i}} are random. Random ε i {\displaystyle \varepsilon_{i}} are called turbulence, noise or simply errors (later in the article they will conflict with residuals; see errors and debris in statistics). Note that to include a constant in the model above, you can choose to introduce the constant as variable β K+1 {\displaystyle \beta_{K+1}} with the latest newly introduced column X Unity, for example, X i (K+1) = 1 {\displaystyle X_{i(K+1)}=1} for all i {\displaystyle i}. Note that although y i, {\displaystyle y_{i},} are visible as sample responses, the following statements and arguments include assumptions, proofs and others assume that under the only condition of knowing X i j, {\displaystyle X_{ij},} but not y i. {\displaystyle y_{i}.} Gauss–Markov's assumptions relate to the set of random error variables, ε i {\displaystyle \varepsilon_{i}}: they mean zero: E ⁡ [ε i] = 0. {\displaystyle \operatorname {E} [\varepsilon _{i}]=0.} They are hemuscodistic, all of which have the same limited variance: Var ⁡ (ε i) = σ 2 < ∞ {\displaystyle \operatorname {Var} (\varepsilon_{i})=\sigma^{2}<\infty} for all i {\displaystyle i} and distinct error terms are uncorrelated: Cov (ε i, ε j) = 0, ∀ i ≠ j. {\displaystyle {\text{Cov}}(\varepsilon _{i},\varepsilon _{j})=0,\forall ieq j.} linear estimator β j {\displaystyle \beta_{j}} a linear compound β ^j = c 1 j y 1 + ⋯ + c n j y n {\displaystyle {\widehat {\beta}}_{j}=c_{1j}y_{1}+\cdots+c_ c_{1j}y_{1}+\cdots+c_ <0>{nj}y_ {n}} where coefficients c i j {\displaystyle c_{ij}} are not allowed to depend on the underlying coefficients β j {\displaystyle \beta_{j}}, since they are not visible, but are not allowed to depend on the values X i j {\displaystyle X_ are {ij}}, since this data is visible. (The dependence of coefficients on any X i j {\displaystyle y_ X_{ij}} ε is normally nonlinear; An impartial estimator is said to be if and only if E ⁡ [β^j] = β j {\displaystyle \operatorname {E} \left[{\widehat {\beta}}_{j}\right]=\beta_{j}} regardless of the X values i j {\displaystyle X_{ij}}. Now let ∑ j = 1 K λ j β j {\displaystyle \sum olimits _{j=1}^{K}\lambda_{j}\beta_{j}} is some linear combination of coefficients. Then the mean square error of the corresponding estimate E ⁡ [∑ j = 1 K λ j (β ^j − β j) 2] , {\displaystyle \operatorname {E} \left[\left(\sum_{j=1 {K}\lambda_{j}\left({\widehat{\beta}}_{j}-\beta_{j}\right)\right)^{2}\right],} in other words, expect a weighted sum square (across parameters) from بین برآورد کننده ها و پارامترهای مربوطه تخمین زده شود. (از آنجا که ما در حال در نظر گرفتن مورد که در آن تمام برآورد پارامتر بی طرفانه هستند، این خطای مربع متوسط همان واریانس ترکیب خطی است.) بهترین برآورد کننده خطی بی طرفانه (BLUE) بردار β {\displaystyle \beta } از پارامترها β j {\displaystyle \beta _{j}} یکی با کوچکترین خطای مربع متوسط برای هر بردار λ {\displaystyle \lambda } از پارامترهای ترکیبی خطی است. این معادل شرطی است که Var ⁡ (β ~) − Var ⁡ (β ^) {\displaystyle \operatorname {Var} \left({\widetilde {\beta }}\right)-\operatorname {Var} \left({\widehat {\beta }}\right)} is a positive semi-definite matrix for every other linear unbiased estimator β ~ {\displaystyle {\widetilde {\beta }}} . برآورد کننده معمولی حداقل مربع (OLS) تابع β ^ = (X ′ X) − 1 X ′ y {\displaystyle {\widehat {\beta }}=(X'X)^{-1}X'y} of y {\displaystyle y} and X {\displaystyle X} (where X ′ {\displaystyle X'} de transpose of X {\displaystyle X}) that minimizes the sum of squares of residuals (misprediction amounts): ∑ i = 1 n (y i − y ^ i) 2 = ∑ i = 1 n (y i − ∑ j = 1 K β ^ j X i j) 2 . {\displaystyle \sum _{i=1}^{n}\left(y_{i}-{\widehat {y}_{i}\right)^{2}=\sum _{i=1}{n}\left(y_{i}-\sum _{j=1}^{K}{\widehat {\beta }}_{j}X_{ij}\right)^{2}.} این تئوری اکنون بیان می کند که برآورد کننده OLS یک BLUE است. ایده اصلی اثبات این است که برآورد کننده کمترین مربعات با هر برآورد کننده بی طرفانه خطی صفر همسو نیست، i.e. , with every linear combination a 1 y 1 + ⋯ + a n y n {\displaystyle a_{1}y_{1}+\cdots +a_{n}y_{n}} whose coeffients do not depend upon the unobservable β {\displaystyle \beta } but whose expected value is always zero. در اثبات اظهار نظر که OLS واقع به حداقل رساندن مجموع مربع از باقی مانده ممکن است به عنوان زیر با محاسبه ماتریس Hessian ادامه و نشان می دهد که آن را مثبت قطعی است. تابع MSE است f که می خواهیم به حداقل برسونیم (β 0 , β 1 , … , β p) = ∑ i = 1 n (y i − β 0 − β 1 x i 1 − ⋯ − β p x i p) 2 {\displaystyle f(\beta _{0},\beta {1},\dots ,\beta _{p})=\sum _{i=1}^{n}(y_{i}-\beta _{0}-\beta _{1}x_{i1}-\dots -\beta _{p}x_{ip})^{2}} for a multiple regression model with p variables. مشتق اول d d β → f = − 2 X T (y → − X β →) = − 2 [∑ i = 1 n (y i − ⋯ − β p x i p) ∑ i = 1 n x i 1 (y i ⋯ − β p x i p) ⋮ ∑ i = 1 n x i p (y i − ⋯ − β p x i p)] = 0 → p + 1 {\displaystyle {\begin{aligned}{\frac {d}{d{d{\ overrightarrow {\beta }}} f&=-2X^{T}({\overrightarrow {y}}-X{\overrightarrow {\beta }}})\\&=-2{\begin{bmatrix}\sum _{i=1}^{n}(y_{i}-\dots -\beta _{p}y_ x_{ip})\\\sum _{i=1}^{n}x_{i1}(y_{i}-\dots -\beta _{p}x_{ip})\\\vdots \\\sum _{i=1}^{n}x_{ip}(y_{i}-\dots -\beta {0}} _{p+1}\end{aligned}}} ,where X is the design matrix X = [1 x 11 … x 1 p 1 x 21 … x 2 p ... 1 x n 1 … x n p] ∈ R n × (p + 1) ; n ⩾ p + 1 {\displaystyle X={\begin{bmatrix}1&x_{11}&\dots &x_{1p}\\1&x_{21}&\dots &x_{2p}\\&&\dots \\1&x_{n1}&\\dots &x_{np}\end{bmatrix}}\in \mathbb {R} ^{n\times (p+1)};\qquad n\geqslant p+1} The Hessian matrix of second derivatives is H = 2 [n ∑ i = 1 n x i 1 … ∑ i = 1 n x i p ∑ i = 1 n x i 1 ∑ i = 1 n x i 1 2 … ∑ i = 1 n x i 1 x i p ⋮ ⋮ ⋱ ⋮ ∑ i = 1 n x i p ∑ i = 1 n x i p x i 1 … ∑ i = 1 n x i p 2] = 2 X T X {\displaystyle {\mathcal {H}}=2{\begin{bmatrix}n&\sum _{i=1}^{n}x_{i1}&\dots &\sum _{i=1}^{n}x_{ip}\\\sum _{i=1}^{n}x_{i1}&\sum _{i=1}^{n}x_{i1}} ^{2}&\dots &\sum _{i=1}^{n}x_{i1}x_{ip}\\\vdots &\vdots &\ddots &\vdots \\\sum _{i=1}. {n}x_{ip}&\sum _{i=1}^{n}x_{ip}x_{i1}&\dots &\sum _{i=1}^{n}x_{ip}^{2}\end{bmatrix}}=2X^{T}X} Assuming the columns of X {\displaystyle X} are linearly independent so that X T X {\displaystyle X^{T}X} is invertible, let X = [v 1 → v 2 → … v → p + 1] {\displaystyle X={\begin{bmatrix}{\overrightarrow {v_{1}}}&{\overrightarrow {v_{2}}}&\dots &{\overrightarrow {v}}_{p+1}\end{bmatrix}}} , then k 1 v 1 → + ⋯ + k p + 1 v → p + 1 = 1 = 1 ⟺ k 1 = ⋯ = k p + 1 = 0 {\displaystyle k_{1}{\overrightarrow {v_{1}}}}+\dots +k_{p+1}} {\overrightarrow {v}}_{p+1}=0\iff k_{1}=\dots =k_{p+1}=0} Now let k → = (k 1 , … , k p + 1) T ∈ R (p + 1) × 1 {\displaystyle {\overrightarrow {k}}=(k_{1},\dots ,k_{p+1})^{T}\in \mathbb {R} ^{(p+1)\times 1}} be an eigenvector of H {\displaystyle {\mathcal {H}}} . k → ≠ 0 → ⟹ (k 1 v 1 → + ⋯ + k p + 1 v → p + 1) 2 > 0 {\displaystyle {\overrightarrow {k}}eq {\overrightarrow {0}}\implies (k_{1}{\overrightarrow {v_{1}}}}+\dots +k_{p+1}{\overrightarrow {v}__{p+1})^{2}>0} in terms of vector multiplication, یعنی [k 1 … k p + 1] [v 1 → ⋮ v → p + 1] [v 1 → … v → p + 1] [k 1 ⋮ k p + 1] = k → T H k → = λ k → T k → > 0 {\displaystyle {\begin{bmatrix}k_{1}&\dots &k_{p+1}\end{bmatrix}}{\begin{bmatrix}{\overrightarrow {v_{1}}}\\\vdots \\{\overrightarrow {v}}_{p+1}\end{bmatrix}}{\begin{bmatrix}{\overrightarrow {v_{1}}}&\dots &{\overrightarrow {v}_{p+1}\end{bmatrix}}{\begin{bmatrix}k_{1}\\\vdots \\k_ {p+1}\end{bmatrix}}={\overrightarrow {k}}^{T}{\mathcal {H}}{\ overrightarrow {k}}=\lambda {\overrightarrow {k}}^{T}{\overrightarrow {k}}>0} where λ {\displaystyle \lambda } is the eigenvalue corresponding to k → {\displaystyle {\overrightarrow {k}}} . همچنین k → T k → = ∑ i = 1 p + 1 k i 2 > 0 ⟹ λ > 0 {\displaystyle {\overrightarrow {k}^{T}{\overrightarrow {k}}=\sum _{i=1}^{p+1}k_{i}^{2}>0\implies \lambda >0} Finally, as eigenvector k → {\displaystyle {\overrightarrow {k}}} was arbitrary, it means eigenvalues of H {\displaystyle {\mathcal {H}}} are بنابراین H {\displaystyle {\mathcal {H}}} قطعی مثبت است. بنابراین β → = (X T X) − 1 X T Y {\displaystyle {\overrightarrow {\beta }}=(X^{T}X)^{-1}X^{T}Y} در واقع یک حداقل محلی است. Proof Let β ~ = C y {\displaystyle {\tilde {\beta }}=Cy} be another linear estimator of β {\displaystyle \beta } with C = (X ′ X) − 1 X ′ + D {\displaystyle C=(X'X)^{-1}X'+D} where D {\displaystyle D} is a K × n {\displaystyle K\times n} non-zero matrix. بنابراین هدف این است که نشان دهد که چنین برآوردگر دارای واریانس کوچکتر از β ^ ، {\displaystyle {\widehat {\beta }},} همانطور که ما در حال محدود کردن به برآوردکننده های بی طرفانه، حداقل میانگین خطای مربعی نشان می دهد حداقل واریانس. برآوردگر OLS است. محاسبه می کنیم: E ⁡ [β ~] = E ⁡ [C y] = E ⁡ [((X ′ X) − 1 X ′ + D) (X β + ε)] = ((X ′ X) − 1 X ′ + D) X β + ((X ′ X) − 1 X ′ + D) E ⁡ [ε] = ((X ′ X) − 1 X ′ + D) X β E ⁡ [ε] = 0 = (X ′ X) − 1 X ′ X β + D X β = (I K + D X) β . {\displaystyle {\begin{aligned}\operatorname {E} \left[{\tilde {\beta }}\right]&=\operatorname {E} [Cy]\\&=\operatorname {E} \left[\left((X'' X)^{-1}X'+D\right)(X\beta +\varepsilon)\right]\\&=\left((X'X)^{-1}X'+D\right)X\beta +\left((X'' X)^{-1}X'+D\right)\operatorname {E} [\varepsilon]\\&=\left((X'X)^{-1}X'+D\right)X\beta &&\operatorname {E} [\\\\ operatorname [\\\ varepsilon]=0\\&=(X'X)^{-1}X'X\beta +DX\beta \\&=(I_{K}+DX)\beta .\\\end{aligned}}} بنابراین، از آنجا که β {\displaystyle \beta } غیر قابل تحمل است، β ~ {\displaystyle {\tilde {\beta }} بی طرفانه است اگر و تنها اگر D X = 0 {\displaystyle DX=0} . سپس: Var ⁡ (β ~) = Var ⁡ (C y) = C Var (y) C ′ = σ 2 C C ′ = σ 2 ((X ′ X) − 1 X ′ + D) (X (X ′ X) − 1 + D ′) = σ 2 ((X ′ X) − 1 X ′ X (X ′ X) − 1 + (X ′ X) − 1 X ′ D ′ + D X (X ′ X) − 1 + D D ′) = σ 2 (X ′ X) − 1 + σ 2 (X ′ X) − 1 (D X) ′ + σ 2 D X (X ′ X) − 1 + σ 2 D D ′ = σ 2 (X ′ X) − 1 + σ 2 D D ′ D X = 0 = Var ⁡ (β ^) + σ 2 D D ′ σ 2 (X ′ X) − 1 = Var ⁡ (β ^) {\displaystyle {\begin{aligned}\operatorname {Var} \left({\tilde {\beta }}\right)&=\operatorname {Var} (Cy)\\&=C{\text{ Var}}(y)C'\\&=\sigma ^{2}CC'\\&=\sigma ^{2}\left((X'X)^{-1}X'+ D\right)\left(X(X'X)^{-1}+D'\right)\\&=\sigma ^{2}\left((X'X)^{-1}X'X(X'X)^{-1}+(X'X)^{-1}X'D'+DX(X'X)^{-1}+DD'\right)\\&=\sigma ^{2}(X'X)^{- 1}+\sigma ^{2}(X'X)^{-1}(DX)'+\sigma ^{2}DX(X'X)^{-1}+\sigma ^{2}DD'\\&=\sigma ^{2}(X'X)^{-1}+\sigma ^{2}DD'&&DX=0 \\\\operatorname {Var} \left({\widehat {\beta }}\right)+\sigma ^{2}DD'&&\sigma ^{2}(X'X)^{-1}=\operatorname {Var} \left({\widehat {\beta }}\right)\end{aligned}}} Since DD' is a positive semidefinite matrix, ⁡ (β ~) {\displaystyle \operatorname {Var} \left({\tilde {\beta }}\right)} exceeds Var ⁡ (β ^) {\displaystyle \operatorname {Var} {\beta }} \right)} by a positive semidefinite matrix. اظهارات در مورد اثبات آن را به عنوان قبلا بیان شده است، شرط Var ⁡ (β ~) − Var ⁡ (β ^) {\displaystyle \operatorname {Var} \left({\tilde {\beta }}\right)-\operatorname {Var} \left ({\widehat {\beta }}\right)} is a positive semidefinite matrix معادل با اموالی است که بهترین برآورد کننده خطی بی طرفانه l t β {\displaystyle \ell ^{t}\beta } l t β ^ {\displaystyle \ell ^{t}{\widehat {\beta }}} بهترین به این معنا که حداقل واریانس دارد). برای دیدن این، اجازه دهید) l t β ~ {\displaystyle \ell ^{ t}{\tilde {\beta }}} یکی دیگر از برآورد کننده های خطی بی طرفانه l t β {\displaystyle \ell ^{t}\beta } . Var ⁡ (l t β ~) = l t Var ⁡ (β ~) l = σ 2 l t (X ′ X) − 1 l + l t D D t l = Var ⁡ (l t β ^) + (D t l) t (D t l) σ 2 l t (X ′ X) − 1 l = Var ⁡ (l t β ^) = Var ⁡ (l t β ^) + ‖ D t l ‖ ⩾ Var ⁡ (l t β ^) {\displaystyle {\begin{aligned}\operatorname {Var} \left(\ell ^{t}{\tilde {\beta }}\right)&=\ell ^{t}\operatorname {Var} \left({\tilde {\beta }}\right)\ ell \\&=\ \sigma ^{2}\ell ^{t}(X'X)^{-1}\ell +\ell ^{t}DD^{t}\ell \\&=\operatorname {Var} \left(\ell ^{t}{\widehat {\beta }}\right)+(D^{t}\ell)^{t}(D^{t}\ell)&\sigma ^{2}\ell ^{t}(X'X)^{-1}\ell =\operatorname {Var} \left(\ell ^{t}{\widehat {\beta }}\right)\\&=\operatorname {Var} \left(\ell ^{t}{\widehat {\beta }}\right)+\| D^{t}\ell \|\\&\geqslant \operatorname {Var} \left(\ell ^{t}{\widehat {\beta }}\right)\end{aligned}}} Moreover, equality holds if and only if D t l = 0 {\displaystyle D^{t}\ell =0} . ما محاسبه l t β ~ = l t (((X ′ X) − 1 X ′ + D) Y) از بالا = l t (X ′ X) − 1 X ′ Y + l t D Y = l t β ^ (D t l) t Y = l t β ^ D t l = 0 {\displaystyle {\begin{aligned}\ell ^{t}{\tilde {\beta }}&=\ell ^{t}\left(((X'X)^{-1}X'+D)Y\right)&&{\text{ from above}}\\&=\ell ^{t}(X'X)^{-1}X 'Y+\ell ^{t}DY\\&=\ell ^{t}{\widehat {\beta }}+(D^{t}\ell)^{t}Y\\&=\ell ^{ t}{\widehat {\beta }}&&D^{t}\ell =0\end{aligned}}} This proves that the equality holds if and only if l t β ~ = l t β ^ {\displaystyle \ell ^{t}{\tilde {\beta }}=\ell ^{t}{\widehat {\beta }}} which gives the uniqueness of the OLS estimator as a BLUE. برآوردگر حداقل مربعات عمومی شده حداقل مربعات (GLS)، توسعه یافته توسط Aitken,[5] را گسترش می دهد گاوس–مارکوف به مورد که در آن بردار خطا دارای یک X {\displaystyle \mathbf {X} } رگرسیونگرها (پارامترهای مورد علاقه) در ماتریس طراحی OLS، نیز آبی است. تئوری گاوس-مارکوف همانگونه که در اقتصادسنجی بیان شده است در بیشتر درمان های Aitken ماتریس کوواریانس غیر اسکالر است. [6] برآورد کننده در نمونه های مکرر ثابت فرض می شوند. این فرض برای یک علم عمدتاً غیر تجربه ای مانند اقتصادسنجی نامناسب تلقی می شود. [7] در عوض، فرضیات { theorem are stated conditional on X {\displaystyle \mathbf {X} } . Linearity of dependent variable is assumed to be a linear function of the variables specified in the model. Specifications must be linear in its parameters. This does not mean that there should be a linear relationship between independent and dependent variables. Independent variables can take nonlinear shapes as long as the parameters are linear. equation y = β 0 + β 1 x 2 , {\displaystyle y=\beta_{0}+\beta_{1}x^{2},} qualified as linear while y = β 0 + β 1 x {\displaystyle y=\beta_{0}+\beta_{1}^{2}x} can be transformed to be linear by replacing β 1 2 {\displaystyle\beta_{1}^{2}} by another parameter, γ {\displaystyle \gamma} . An equation with a parameter dependent on an independent variable does not qualify as linear, for example y = β 0 + β 1 (x) ¢ x {\displaystyle y=\beta_{0}+\\beta _{1}(x)\cdot x} , where β 1 (x) {\displaystyle \beta_{1}(x)} is a function of x {\displaystyle x} . Data conversions are often used to convert an equation into linear form. For example, the Cobb-Douglas function—often used in economics—is nonlinear: Y = A L α K 1 − α e ε {\displaystyle Y=AL^{\alpha}K^{1-\alpha}e^{\varepsilon}} but it can be expressed linearly by capturing the natural logarithm of both sides. :[8] ln ⁡ Y = ln ⁡ A + α ln ⁡ L + (1 − α) ln ⁡ K + ε = β 0 + β 1 ln ⁡ L + β 2 ln ⁡ K + ε {\displaystyle \ln Y=\ln A+\ alpha \ln L+(1-\alpha)\ln K+\varepsilon=\beta_{0}+\beta_{1}\ln L+\beta_{2}\ln K+\varepsilon} covers this assumption of specification issues: assuming that the appropriate functional form is selected and there are no deleted variables. But it should be aware that the parameters that minimized the remnants of the converted equation did not necessarily minimize the remainder of the original equation. The strict exogeneity for all n {\displaystyle n} observations is waiting—subject to regression—of the term error zero:[9] E ⁡ [ε i ∣ X] = E ⁡ [ε i ∣ x 1 , … , x n] = 0. {\displaystyle \operatorname {E} [\,\varepsilon _{i}\mid \mathbf {X}]=\operatorname {E} [\,\varepsilon _{i}\mid \mathbf {x_{1}} ,\dots ,\mathbf {x_{n}}]=0.} where x i = [x i 1 x i 2 … x i k] T {\displaystyle \mathbf {x} _{i}={\begin{bmatrix}x_{i1}&x_{i2}&\dots &;x_{ik}\end{bmatrix}}^{\mathsf {T}}} is the data vector of regressions to view ith, resulting in X = [x 1 T x 2 T … x n T] T {\displaystyle \mathbf {X} ={\begin{bmatrix}\mathbf {x_{1}^{\mathsf {T}}} &\mathbf {x_{2}^{\mathsf {T }}} &\dots &\mathbf {x_{n}^{\mathsf {T}}} \end{bmatrix}}^{\mathsf {T}}} is the data matrix or design matrix. Geometrically, this assumption implies that x i {\displaystyle \mathbf {x} _{i}} and ε i {\displaystyle \varepsilon_{i}} are symmetrical, so that their inner (As one of their crossover moments) is zero. E ⁡ [x j × ε i] = [E ⁡ [x j 1 × ε i] E ⁡ [x j 2 × ε i] ⋮ ⁡ [x j k ¢ ε i]] = 0 for all i , j ∈ n {\displaystyle \operatorname {E} [\,\mathbf {x} _{j}. \cdot \varepsilon _{i}\,]={\begin{bmatrix}\operatorname {E} [\,{x}_{j1}\cdot \varepsilon _{i}\,]\\\\\operatorname {E} [\,{x}_{j2}\cdot \varepsilon _{i}\,]\\\vdots \\\operatorname {E} [\ [\ ,{x}_{jk}\cdot \varepsilon _{i}\ ,]\end{bmatrix}}=\mathbf {0} \quad {\text{for all}}i,j\in n} This assumption is violated if explanation variables, for example, when measured by error, or endogenous. [10] Endogenousity can be the result of simulation, where the existence between both dependent and independent variables flows back and forth. Variable instrumental techniques are commonly used to address this problem. The full rank of the sample data matrix X {\displaystyle \mathbf {X} } must have the full column rank. rank ⁡ (X) = k {\displaystyle \operatorname {rank} (\mathbf {X})=k} Otherwise X' X {\displaystyle \mathbf {X'X} } is not invertible and the OLS estimator cannot be calculated. The violation of this multicolinari assumption is complete, meaning that some explanation variables are linearly dependent. A scenario in which this will happen is called a dummy variable trap, when a baseline dummy variable is not eliminated, resulting in a complete correlation between the fabricated variables and the constant term. [11] Multicollinearity (as long as it's not complete) can now result in less efficient, but still estimates impartially. Estimates will be less accurate and highly sensitive than certain sets of data. [12] Multicollinearity can be detected from the conditions number or the factor of variance swelling, among other tests. Spherical errors of the product out of the error vector must be spherical. E ⁡ [ε ε T ∣ X] = Var ⁡ [ε ∣ X] = [σ 2 0 … 0 0 σ 2 … 0 ⋮ ⋮ ⋱ ⋮ 0 0 … σ 2] = σ 2 I with σ 2 > 0 {\displaystyle \operatorname {E} [\,{\boldsymbol {\varepsilon }}{\boldsymbol {\varepsilon ^{\\\\\ mathsf {T}}}}\mid \mathbf {X}]=\operatorname {Var} [\,{\boldsymbol {\varepsilon }}\mid \mathbf {X}]={\begin{bmatrix}\sigma ^{2} &0&\dots &0\0&\sigma ^{2}&\dots &0\\\vdots &\vdots &\ddots &\vdots \\0&0&\dots &\sigma ^{2}\\\\ end{bmatrix}}=\sigma ^{2}\mathbf {I} \quad {\text{with }}\sigma ^{2}>0} This implies the error term has uniform variance (homoscedasticity) and no serial dependence. [13] If this assumption is violated, OLS is still impartial, but inefficient. The term spherical errors will describe the normal multivariate distribution: if Var ⁡ [ε ∣ X] = σ 2 I {\displaystyle \operatorname {Var} [\,{\boldsymbol {\varepsilon}}\mid\mathbf {X}]=\sigma^{2}\mathbf {I}} in the multivariate normal density, Then the equation f (ε) = c {\displaystyle f(\varepsilon)=c} is the ball-driven formula μ The radius σ is the next nth space. [14] Heteroskedasticity occurs when the error value is in line with an independent variable. For example, in a regression on food expenditures and income, error is correlated with income. Low-income people generally spend the same amount on food, while high-income people may spend quite a lot or as much as low-income people. Heteroskdastic can also be caused by changes in measurement methods. For example, as statistical offices improve their data, the measurement error decreases, so the term error decreases over time. This assumption is violated when there is a clause itself. If adjacent observations are also above the mounted regression line, it is possible to visualize on a data scheme where a given observation is more likely to lie above an installed line. Fusion itself is common in time series data where a data series may experience laxity. If an dependent variable takes some time to absorb a shock completely. Spatial self-life can also occur, geographic regions are most likely to have similar errors. Self-conflict may be the result of misinformation, such as choosing the wrong functional form. In these cases, correcting the specifications is one of the possible ways to deal with the co-ording itself. In the presence of spherical errors, the estimator of the least publiced squares can be shown to be blue. [6] See also independent and identical distribution of random linear regression variables measuring the uncertainty of other impartial statistics best linear impartial prediction (BLUP) minimum variance impartial estimator (MVUE) sources ^ see Chapter 7 of Johnson, R.A.; Wichern, D.W. (2002). Applied multivariate statistical analysis. 5. Prentice Hall. ^ Theil, Henri (1971). Best Linear Unbiased Estimation and Prediction. Econometric principles. New York: John Wiley & Sons. pp. 119–124. ISBN 0-471-85845-5. ^ Plackett, R. L. (1949). A Historical Note on the Method of Least Squares. biometrics . 36 (3/4): 458–460. doi:10.2307/2332682. ^ David, F. N.; Neyman, J. (1938). Extension of the Markoff theorem on least squares. Memoirs of Statistical Research. 2: 105–116. OCLC 4025782. ^ a b Aitken, A.C. (1935). On Least Squares and Linear Combinations of Observations. Proceedings of the Royal Society of Edinburgh. 55: 42–48. doi:10.1017/S0370164600014346. ^ a b Huang, David S. (1970). Regression and econometric methods. New York: John Wiley & Sons. Page 127–147. ISBN 0-471-41754-8. ^ Hayashi, Fumio (2000). Econometrics. Princeton University Press. p. 13. ISBN 0-691-01018-8. ^ Walters, A. A. (1970). Introduction to Econometrics. New York: W. W. Norton. p. 275. ISBN 0-393-09931-8. ^ Hayashi, Fumio (2000). Econometrics. Princeton University Press. p. 7. ISBN 0-691-01018-8. ^ Johnston, John (1972). Econometric Methods (Second ed.). New York: McGraw-Hill. pp. 267–291. ISBN 0-07-032679-7. ^ Jeffrey (2012). Introductory Econometrics (Fifth international ed.). southwest . p. 220. ISBN 978-1-111-53439-4. ^ Johnston, John (1972). Econometric Methods (Second ed.). New York: McGraw-Hill. Page 159–168. ISBN 0-07-032679-7. ^ Hayashi, Fumio (2000). Econometrics. Princeton University Press. p. 10. ISBN 0-691-01018-8. ^ Ramanathan, Ramu (1993). Nonspherical Disturbances. Statistical methods in econometrics. Academic Press. Page 330–351. ISBN 0-12-576830-3. Continue reading Davidson, James (2000). Statistical analysis of regression model. Econometric theory. Oxford: Blackwell. Page 17–36. ISBN 0-631-17837-6. Goldberger, Arthur (1991). Classical Regression. A course in econometrics. Cambridge: Harvard University Press. pp. 160–169. ISBN 0-674-17544-1. Theil, Henri (1971). Least Squares and the Standard Linear Model. Econometric principles. New York: John Wiley & Sons. pp. 101–162. ISBN 0-471-85845-5. External links are the first known uses of some mathematical words: G (brief history and name description) proving Gauss Markov's theory for multiple linear regression (causing matrix algebra) proof of Gauss Markov's theory using geometry retrieved from .. wikipedia.org/w/index.php?title=Gauss–Markov_theorem&oldid=993370161