Nonlinear Unbiased Estimation in Linear Regression with Nonnormal Disturbances

Author: Knautz (1999)

Presented by: Gaurav Kumar Rampuria
(BS, Statistics & Data Science, IIT Kanpur)

Outline

- Motivation
- LPQ Estimator Class
- Optimal LPQ Estimator
- Feasible Implementation
- Simulation Design
- Key Findings
- Practical Implications
- Conclusion

Motivation

Key Problem

In the application of the linear regression model there continues to be wide-spread use of the Least Squares Estimator (LSE) due to its theoretical optimality. It is the best unbiased estimator under normality, and the best *linear* unbiased estimator if the normality assumption is dropped. Least Squares (LS) optimal under normality but suboptimal for:

- Skewed error distributions ($m_3 \neq 0$)
- Heavy-tailed distributions (excess kurtosis)

Proposed Solution

- Linear + Quadratic (LPQ) estimators
- The LPQ estimators exploit 3rd and 4th moments (skewness, kurtosis) of the error distribution. Simulation shows the improvement holds even when these moments are estimated from data.

LPQ Estimator Class

Standard linear model:

$$y = X\beta + \epsilon$$

- X: $n \times p$ design matrix (full rank)
- ϵ_i iid with $E(\epsilon_i) = 0$, $Var(\epsilon_i) = \sigma^2$
- Known higher moments: $m_{3i} = E(\epsilon_i^3)/\sigma^3$, $m_{4i} = E(\epsilon_i^4)/\sigma^4$

General form of LPQ Estimator Class:

$$\widetilde{\beta} = X^{-}y + \frac{1}{\sigma}\{y' H_i y\} - \sigma\{\text{tr } H_i\}$$

- H_i, i = 1,...,p are n $\times$ n matrices which are assumed to be symmetric without loss of generality, with the constraint $H_i X = 0$.
- **Unbiasedness**: Maintains $E(\tilde{\beta}_{\text{LPQ}}) = \beta$ through linear+quadratic construction (Expection of the quadratic component is 0)
- Equivariant under translations: $\widetilde{\beta}(y + X\alpha) = \alpha + \widetilde{\beta}(y)$

Optimal LPQ Estimator

Theorem (Optimal LPQ Estimator)

Minimizes matrix MSE:

$$\widetilde{\beta}_{LPQ} = X^{+}y + D\widetilde{K}\left\{z * \mu - \frac{1}{\sigma}(M * M)^{-1}(z * z) + \sigma 1_n\right\}$$

where $z = My$ (LS residuals), $M = I - XX^{+}$

Covariance Structure

The LPQ estimator dominates least squares whenever the error distribution exhibits asymmetry. Specifically, dominance occurs when $D = X^{+}1_p\mu\prime \neq 0$ which happens precisely when $\mu \neq 0$. The dispersion matrix of the optimal LPQ estimator is:

$$\mathrm{Cov}(\widetilde{\beta}_{LPQ}) = \sigma^2\left\{(X'X)^{-1} - D\widetilde{K}D'\right\}$$

Feasible Implementation

- Estimation of the unknown moments is made on the basis of the vector of LS residuals $z = My$. $\hat{\sigma}^2_{LS} = \frac{z'z}{\nu}$, where $\nu = n - p$.
- Estimators for m_3 and $m_4 - 3$ are derived using:

$$\tilde{m}_3 = \frac{(1/n)\sum_{i=1}^{n}(z_i - \bar{z})^3}{s_z^3}, \tilde{m}_{43} = \frac{(1/n)\sum_{i=1}^{n}(z_i - \bar{z})^4}{s_z^4} - 3, \quad (1)$$

where $s_z^2 = \frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{z})^2$, and $m_3 \leq \tilde{m}_{43} + 2$.

Consistentcy criteria:

1. $\lim_{n\to\infty} \frac{1}{n} X'X = Q > 0$
2. $E|\epsilon_i|^{4+\delta} < \infty$ for some $\delta > 0$
3. $\lambda_{\min}(M * M) = c > 0$

Practical Computation

$$\widetilde{\beta}^*_{LPQ} = X^+ y + \hat{m}_3 X^+ \widetilde{K}^* \{\hat{m}_3 z - \frac{1}{\hat{\sigma}}(M * M)^{-1}(z * z) + \hat{\sigma} 1_n\}$$

Simulation Design

Dataset: A realistic design matrix derived from a fuel consumption data set that covers 48 US states, with variables that include tax rates, driver's license percentages, average income, and highway lengths was considered for our Monte Carlo simulation.

Distribution	Skewness (m_3)	κ	MSE Ratio (LPQ/LS)
Normal (Norm)	0.00	0.00	1.00-1.03
Laplace (Lap)	0.00	3.00	0.95-0.98
Lognormal (LogN)	6.19	110.94	0.26-0.29
Exponential (Exp)	2.00	6.00	0.49-0.53
Pareto ($a = 5$)	4.65	70.80	0.32-0.36
Pareto ($a = 10$)	2.81	14.83	0.40-0.44
Beta (Beta)	-2.66	6.57	0.23-0.24
Gamma (Gam)	2.83	12.00	0.32-0.35

Key Findings

The simulation results strongly support the theoretical predictions. For asymmetric distributions, the LPQ estimator demonstrates substantial improvements over least squares:

- **25-50% MSE reduction** for asymmetric distributions
- **65-75% probability** of LPQ estimates being closer to the true parameters than LS estimates
- Exhibits smaller empirical bias than LS for nonsymmetric distributions
- *Coverage probability*: LPQ-based bootstrap CIs achieve near-nominal coverage (94–97%), **better than LS intervals, for skewed errors**.
- *Interval length*: LPQ intervals are on average shorter (ratio < 1) than LS intervals; often **70–90%** of LS length.

Symmetric distributions (Normal and Laplace):

MSE ratios near unity, indicating comparable performance. This validates that the LPQ approach does not sacrifice performance under ideal conditions while providing substantial gains when assumptions are violated.

Practical Conisderations

LPQ estimation should be seriously considered when:

- Error distributions exhibit skewness
- Sample sizes are moderate to large ($n \geq 50$) (the consistency results require asymptotic conditions)
- Precision gains are valuable relative to computational costs
- Robust inference needed for non-normal data

Limitations

- Requires iid asymmetric errors
- Assumes finite fourth moments
- Computationally intensive vs LS

Conclusion

- The paper introduces us to the LPQ class of non linear estimator demonstrates how they outperform LS estimators under non normal and asymmetric error distributions. These LPQ estimators are **unbiased, equivariant, and consistent** estimators
- Using a simulation study we clearly see how they outpeform LS estimators on various metrics including MSE Ratios, Bias ratio, and Interval lengths.
- These results can find extensive use in financial and economic applications where the error distributions are usually non normal and asymmetric.

Future Work

- Extension to heteroskedastic models
- Comparison with Stein estimators
- Robustness to moment estimation errors