Bayesian Non-linear Quantile Regression with Application in Decline Curve Analysis for Petroleum Reservoirs.

Abstract

by

YOUJUN LI

Quantile regression (QR) approach, proposed by Koenker and Bassett (1978) [1], is being widely used in statistical literature for linear regression models as an alternative to the conventional mean regression (MR) model. In quantile regression models the quantiles of the conditional distribution of the response variable are expressed as functions of the covariates. The quantile regression is particularly useful when the conditional distribution is heterogeneous and does not have a standard shape, such as an asymmetric, fat-tailed, or truncated distribution. Compared to conventional mean regression (MR), quantile regression (QR) is more robust to outliers and misspecification of the error distribution. Koenker (2000)[2], Koenker and Hallock (2001)[3] and Koenker (2005)[4] provide a thorough review and study of the quantile regression method.

From a Bayesian point of view, Kottas & Gelfand (2001)[5] considered median regression, which is a special case of quantile regression, and discussed non-parametric modeling for the error distribution based on either Plya tree or Dirichlet process priors. Regarding general quantile regression, Yu & Moyeed (2001)[6] proposed a Bayesian modeling approach by using the asymmetric Laplace distribution for the error distribution. Kottas & Krnjajic (2009)[7] developed Bayesian semi-parametric models for quantile regression using Dirichlet process mixtures for the error distribution.

All these Bayesian approaches considered the quantile regression model in a linear regression model or generalized linear regression model setup. In this thesis, we consider non-linear quantile regression model where the quantiles of the conditional distribution of the response variable are expressed as non-linear functions of the covariates. The non-linear quantile regression method would be used in a real field data related to production history in hydrocarbon reservoirs. Analysis of declining production rates and forecasting future performance of oil and gas wells is called decline curve analysis in petroleum engineering literature. Decline curve analysis is a basic tool for estimating recoverable reserves which can be used when the production history is long enough that a decline trend can be identified. Decline curve analysis is mainly a nonlinear regression model where certain types of standard curves are fit based on the past production performance and is then extrapolated to predict the future well performance. Arps (1945)[8] provided some theoretical justification based on physics about those standard types of curves, viz. hyperbolic, harmonic and exponential. The conventional way of decline curve analysis in the petroleum engineering community is to fit a mean regression model using least square method on the transformed production data. The fitted curve is defined as the P50 curve which is extrapolated for future prediction and reported to security exchange commission (SEC) for the reserves estimates. Surprisingly the definition of the P50 estimate in the SEC handbook is actually the 50-th percentile, not the predicted mean. When using the least square estimate what is being reported is the predicted mean of the curve, not the 50-th percentile or P50. Similarly, the prediction interval is used to report the P10 and P90 (the 10th percentile and the 90th percentile respectively) reserves. The use of quantile regression instead of the conventional mean regression would be appropriate in this context as the fitted quantile regression curves would have the correct interpretation for the predicted reserves. Here we could consider the hyperbolic curve as the non-linear quantile regression function when the exponential and the harmonic curve can be viewed as a special case of the hyperbolic curve. The information from the neighboring well production history and similar type of reservoirs would be incorporated as priors for the three decline parameters in the hyperbolic curve. The asymmetric Laplace distribution would be used for modeling the likelihood. The posterior distribution of the regression coefficients and other parameters is intractable mainly due to the non-linearity in the quantile regression function, hence a combination of Gibbs sampling and Metropolis Hastings algorithm would be used to sample from the posterior. A quantitative assessment of the uncertainty of the decline parameters and the future prediction would be provided.

Bibliography

[1] Roger Koenker and Gilbert Bassett Jr. Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50, 1978.

[2] Roger Koenker. Galton, edgeworth, frisch, and prospects for quantile regression in econometrics. *Journal of Econometrics*, 95(2):347–374, 2000.

[3] Roger Koenker and Kevin Hallock. Quantile regression: An introduction.

[4] Roger Koenker. *Quantile regression*. Number 38. Cambridge university press, 2005.

[5] Athanasios Kottas and Alan E Gelfand. Bayesian semiparametric median regression modeling. *Journal of the American Statistical Association*, 96(456):1458–1468, 2001.

[6] Keming Yu and Rana A Moyeed. Bayesian quantile regression. *Statistics & Probability Letters*, 54(4):437–447, 2001.

[7] Athanasios Kottas and MILOVAN KRNJAJIĆ. Bayesian semiparametric modelling in quantile regression. *Scandinavian Journal of Statistics*, 36(2):297–319, 2009.

[8] Jan J Arps et al. Analysis of decline curves. *Transactions of the AIME*, 160(01):228–247, 1945.

[9] James L Powell. Censored regression quantiles. *Journal of econometrics*, 32(1):143–155, 1986.

[10] Roger Koenker and Jose AF Machado. Goodness of fit and related inference processes for quantile regression. *Journal of the american statistical association*, 94(448):1296–1310, 1999.

[11] Moshe Buchinsky. Changes in the us wage structure 1963-1987: Application of quantile regression. *Econometrica: Journal of the Econometric Society*, pages 405–458, 1994.

[12] Victor Chernozhukov and Christian Hansen. The effects of 401 (k) participation on the wealth distribution: an instrumental quantile regression analysis. *Review of Economics and statistics*, 86(3):735–751, 2004.