

УДК 519.233.5

Восстановление энтропийно-оптимальной плотности при построении линейной регрессионной модели методом максимальной энтропии.

Ю.А. Дубнов^{1,2}

¹Московский физико-технический институт (государственный университет)

²Институт системного анализа РАН (ИСА РАН)

Построение регрессионных моделей является одной из фундаментальных задач статистики и эконометрики. Наиболее известными подходами к оцениванию параметров моделей являются метод наименьших квадратов, метод максимального правдоподобия и байесовский подход. Работа посвящена описанию и экспериментальной апробации альтернативного метода оценивания параметров регрессии - метода максимальной энтропии. В работе приведено описание метода, обсуждены некоторые его модификации и проведено качественное сравнение с наиболее распространенными альтернативами на примере нормальных и смещенных возмущений.

Принцип максимизации энтропии впервые был сформулирован Эдвардом Т. Джейнсом в 1957 году в качестве метода приближенного решения переопределенных систем уравнений. И только в 1996 году А. Голан и др. предложили использовать этот принцип для построения регрессионных моделей в эконометрике и разработали метод максимальной энтропии (GME, Generalized Maximum Entropy) [1].

Как и другие известные методы оценивания, такие как, метод максимального правдоподобия (ММП) и байесовский подход, метод максимальной энтропии (ММЭ) основан на предположении о стохастическом характере модели. Так, параметрами модели полагаются случайные величины, а их вероятностные характеристики восстанавливаются в процессе оценивания. В частности, в ММЭ восстанавливаются энтропийно-оптимальные плотности вероятности параметров модели, реализующие максимум информационной энтропии Больцмана [2], которые далее используются для построения точечных и интервальных оценок параметров.

Таким образом, оценки ММЭ реализуют максимум энтропии рассматриваемой системы, что может трактоваться как наиболее правдоподобные в терминах вероятности оценки при максимальной неопределенности. Такая формулировка обусловлена тем фактом, что при построении ММЭ-оценок не используется никаких дополнительных предположений о системе, кроме априорно заданной информации (например, результаты измерений, закон распределения ошибок и пр.).

Важной особенностью ММЭ является инвариантность относительно закона распределения ошибок. Как правило, при построении регрессионных моделей закон распределения ошибок считается известным, что используется при построении функции правдоподобия в ММП и байесовском подходе. Именно этот факт обеспечивает статистическим методам преимущество по сравнению с МНК в случае, если закон распределения ошибок отличается от традиционного нормального [3]. Здесь ММЭ выгодно отличается от вышеуказанных аналогов в силу отсутствия необходимости задавать закон распределения ошибок в явном виде, а в процессе оценивания будут восстановлены энтропийно-оптимальные плотности не только для параметров модели, но и для шумов.

В работах [4], [5] рассмотрен классический метод максимальной энтропии в дискретном исполнении, авторы приводят результаты экспериментов, демонстрирующие эффективность ММЭ для разного типа ошибок по сравнению с МНК. В этих экспериментах не использовалась априорная информация об ошибках, что эквивалентно равномерному априорному распределению. Однако некоторая модификация ММЭ, рассмотренная в данной работе, предполагает использование априорной информации для повышения точности оценок. В этом случае представляется интересным сравнить результаты различных статистических методов оценивания с одинаковыми начальными условиями.

В данной работе рассмотрена непрерывная реализация ММЭ с априорными распределениями ошибок, заданными по нормальному закону и закону «хи-квадрат», рассмотрен механизм построения энтропийно-оптимальных плотностей и исследована зависимость ММЭ-оценок от объема начальных данных. В ходе численных экспериментов показана состоятельность оценок и эффективность метода при ограниченном объеме входных данных.

Литература

1. *Amos Golan, George G. Judge, Douglas Miller. Maximum Entropy Econometrics: Robust Estimation with Limited Data.* — Chichester, U.K.: John Wiley and Sons Ltd, 1996 — 324p.
2. *Cover T.M., Thomas J.A. Elements of information theory.* — New York: John Wiley and Sons Ltd, 1991. — 561p.
3. *Магнус Я.Р., Катышев П.К., Персецкий А.А. Эконометрика. Начальный курс: Учеб., 6-е изд., перераб. и доп.* — М.: Дело, 2004. — 576с.
4. *Ozan H. Eruygur. Generalized maximum entropy (GME) estimator: formulation and a monte carlo study* — VII. National Symposium on Econometrics and Statistics, Istanbul, Turkey — 2005.
5. *Ximing Wu. A Weighted Generalized Maximum Entropy Estimator with a Data-driven Weight* — Entropy — 2009 — N. 11. 1-x.