The International Conference on Trends and Perspectives in Linear Statistical Inference (LinStat2014)
24 – 28 August 2014 Linköping, Sweden

Nonlinear Inequality Constrained Ridge Regression Estimator

Dr. Huseyin GULER (hguler@cu.edu.tr)
Cigdem KOSAR (ckosar@cu.edu.tr)

Cukurova University,
Department of Econometrics

Presentation Plan

1. MOTIVATION
2. MODEL
3. MULTICOLLINEARITY and the RIDGE REGRESSION ESTIMATOR
4. NONLINEAR INEQUALITY CONSTRAINED RIDGE REGRESSION ESTIMATOR
5. MONTE CARLO EXPERIMENT
6. FINDINGS
7. RESULTS and CONCLUSION

1. Motivation

- Linear Regression Model – Ordinary Least Squares (OLS)
- Inequality Constrained Least Squares (Liew, 1976)
- Nonlinear Inequality Constraints
  - Almost Ideal Demand System – AIDS. Deaton and Muellbauer, 1980
  - Translog Demand System – TL. Jorgenson, Lau and Stoker, 1982
    - Convexity – Concavity, Monotonicty, Nonnegativeness…

- Bayesian Inference Methods (Geweke, 1986)
  - Markov Chain Monte Carlo (MCMC)
    - Metropolis-Hastings algorithm (Metropolis et al, 1953; Hasting, 1970)

- Chua et al (2001) estimates nonlinear inequality constrained AIDS and LogTL models for the US with Bayesian Model Averaging (BMA).

- O’Donnell et al (2001) estimates equality and nonlinear inequality constrained TL demand system for Oklahoma and Teksas with MCMC. They also assume that the model parameters are variable across individuals and time.

- Koop and Potter (2011) estimates a time-varying VAR model for the US between 1953:1-2006:2 with MCMC. They have a restriction that the roots of the VAR polynomial are outside the unit circle.

- Multicollinearity
- Ridge Regression Estimator (RR, Hoerl and Kennard, 1970)
- Inequliaty Constrained Ridge Regression Estimator (Toker vd., 2013)
- As far as we know, multicollinearity problem in a nonlinear inequality constrained regression model is not studied in the literature.
- In this study, we define a Nonlinear Inequality Constrained Ridge Regression Estimator (NICRR).
- Posterior distributions of the parameters are obtained.
- MCMC is suggested to estimate the moments of posterior distributions.
- OLS, RR, Nonlinear Inequality Constrained Least Squares (NICLS) and NICRR are compared with a Monte Carlo study.

2. Model

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1)$$

$$h_r(\boldsymbol{\beta}) \geq 0 \, , r = 1,2, \dots , R \quad (2)$$

- $\boldsymbol{y} = (y_1 \quad y_2 \quad \cdots \quad y_n)'$
- $\boldsymbol{X} = (\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \cdots \quad \boldsymbol{x}_p), \boldsymbol{x}_j = (x_{1j} \quad x_{2j} \quad \cdots \quad x_{nj})'$
- $\boldsymbol{\beta} = (\beta_1 \quad \beta_2 \quad \cdots \quad \beta_p)'$
- $\boldsymbol{\varepsilon} = (\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n)'$
- n: #(observations), p: #(independent variables including a constant), R: #(restrictions).

Assumptions:

- $E(\boldsymbol{\varepsilon}\ /\boldsymbol{X}) = \mathbf{0}$,
- $var(\boldsymbol{\varepsilon}\ /\boldsymbol{X}) = \sigma^2 \boldsymbol{I}_n$,
- $\boldsymbol{X}$ is fixed,
- $rank(\boldsymbol{X}) = p$.

3. Multicollinearity and the Ridge Regression Estimator

- The exact/strong linear relationship between independent variables.
- Exact multicollinearity:
- $a_0 + a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_p\boldsymbol{x}_p = \boldsymbol{0}$
  - $|\boldsymbol{X}'\boldsymbol{X}| = 0$,
  - $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{S}^{-1}\boldsymbol{X}'\boldsymbol{y}$,
  - It is impossible to compute OLS.

- Strong (severe) multicollinearity:
- $a_0 + a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_p\boldsymbol{x}_p \cong \boldsymbol{0}$
  - $|\boldsymbol{X}'\boldsymbol{X}| \cong 0$,
  - OLS can be computed but
  - Magnitudes and signs of the estimates may be different from the expectations,
  - Estimates are highly sensitive to the small changes in the sample,
  - Standard errors become bigger,
  - t statistics become smaller and the null hypotheses are more oftenly accepted,
  - t tests suffer from size distortions,
  - The power of the t tests will drop.

- Ridge regression estimator (Hoerl and Kennard, 1970)

$$\hat{\boldsymbol{\beta}}(k) = \left(\boldsymbol{X}'\boldsymbol{X} + k\boldsymbol{I}_p\right)^{-1}\boldsymbol{X}'\boldsymbol{y} = \boldsymbol{S}_k^{-1}\boldsymbol{X}'\boldsymbol{y},\ k > 0$$

- Adds a positive constant k to the diagonal of $\boldsymbol{X}'\boldsymbol{X}$.

(+): Smaller standard errors.

(+): Stabilized estimates.

(−): A biased estimator.

- For a suitably chosen k, RR is better than OLS in the Mean Square Error (MSE) sense.

- MSE of RR:

$$MSE\left[\widehat{\boldsymbol{\beta}}(k)\right] = cov\left[\widehat{\boldsymbol{\beta}}(k)\right] + bias\left[\widehat{\boldsymbol{\beta}}(k)\right]bias\left[\widehat{\boldsymbol{\beta}}(k)\right]'$$
$$= \sigma^2 \boldsymbol{S}_k^{-1}\boldsymbol{S}\boldsymbol{S}_k^{-1} + k^2\boldsymbol{S}_k^{-1}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{S}_k^{-1}$$

- Hoerl and Kennard (1970) show that RR is better than OLS in the MSE sense for

$$k = \frac{p\sigma^2}{\boldsymbol{\beta}'\boldsymbol{\beta}}$$

- k is estimated by

$$\hat{k}_{HK} = \frac{p\hat{\sigma}^2}{\widehat{\boldsymbol{\beta}}'\widehat{\boldsymbol{\beta}}}$$

- It is possible to obtain RR in a Bayesian framework. Lindley and Smith (1972) show that if the prior information is

$$\boldsymbol{y} \sim N_n(\boldsymbol{X\beta}, \sigma^2 \boldsymbol{I}_n) \quad (3)$$

$$\boldsymbol{\beta} \sim N_p\left(\boldsymbol{0}, \tau^2 \boldsymbol{I}_p\right) \quad (4)$$

then the Bayesian estimate of $\boldsymbol{\beta}$ is

$$\widehat{\boldsymbol{\beta}}^* = \left(\boldsymbol{X}'\boldsymbol{X} + \frac{\sigma^2}{\tau^2}\boldsymbol{I}_p\right)^{-1} \boldsymbol{X}'\boldsymbol{y} \quad (5)$$

If k is replaced with σ^2/τ^2 in (6) then the resulting estimator is RR.

- RR can be also obtained by solving the following minimisation problem:

$$\mathcal{L}(\boldsymbol{\beta}; k) = (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + k\boldsymbol{\beta}'\boldsymbol{\beta} \quad (6)$$

4. Nonlinear Inequality Constrained Ridge Regression Estimator

Model: $$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad (1)$$

Restrictions: $$h_r(\boldsymbol{\beta}) \geq 0 \text{ , } r = 1,2, \dots , R \quad (2)$$

Prior Information: $$\boldsymbol{y} \sim N_n(\boldsymbol{X\beta}, \sigma^2 \boldsymbol{I}_n) \quad (3)$$

$$\boldsymbol{\beta} \sim N_p\left(\mathbf{0}, \tau^2 \boldsymbol{I}_p\right) \quad (4)$$

$$\tau^2 = \frac{\sigma^2}{k} \quad (5)$$

$$f(\boldsymbol{y}/\boldsymbol{\beta}, \sigma^2, \tau^2) = (2\pi\sigma^2)^{-n/2} exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta})\right\} \quad (7)$$

$$f(\boldsymbol{\beta}/\sigma^2, \tau^2) = (2\pi\tau^2)^{-p/2} exp\left\{-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\right\} \quad (8)$$

$$f(\boldsymbol{\beta}, \sigma^2, \tau^2/\boldsymbol{y}) = \frac{f(\boldsymbol{\beta}, \sigma^2, \tau^2, \boldsymbol{y})}{f(\boldsymbol{y})} = \frac{f(\boldsymbol{y}/\boldsymbol{\beta}, \sigma^2, \tau^2) f(\boldsymbol{\beta}, \sigma^2, \tau^2)}{f(\boldsymbol{y})}$$

$$= cf(\boldsymbol{y}/\boldsymbol{\beta}, \sigma^2, \tau^2) f(\boldsymbol{\beta}, \sigma^2, \tau^2) \quad (9)$$

$$f(\boldsymbol{\beta}, \sigma^2, \tau^2) = f(\boldsymbol{\beta}/\sigma^2, \tau^2) f(\sigma^2, \tau^2) I(h) \quad (10)$$

$$I(h) = \begin{cases} 1, \; h_r(\boldsymbol{\beta}) \geq 0 \text{ if } (\forall\, r = 1,2, \dots, R) \\ 0, \; h_r(\boldsymbol{\beta}) < 0 \text{ if (At least for an } r) \end{cases} \quad (11)$$

- When $\tau^2 > 0$ and $\sigma^2 > 0$, the prior distribution of τ^2 and σ^2 is

$$f(\sigma^2, \tau^2) = \frac{1}{\sigma^2}\frac{1}{\tau^2} \tag{12}$$

- If (8) and (12) are replaced in (10)

$$f(\boldsymbol{\beta}, \sigma^2, \tau^2) = (2\pi\tau^2)^{-p/2} exp\left\{-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\right\}\frac{1}{\sigma^2}\frac{1}{\tau^2} I(h) \tag{13}$$

- If (7) and (13) are replaced in (9)

$$f(\boldsymbol{\beta},\sigma^2,\tau^2/\boldsymbol{y}) = cf(\boldsymbol{y}/\boldsymbol{\beta},\sigma^2,\tau^2)f(\boldsymbol{\beta}/\sigma^2,\tau^2)I(h)$$

$$= c(\sigma^2)^{-\frac{n+2}{2}}(\tau^2)^{-\frac{p+2}{2}} \qquad (14)$$

$$\times exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta})-\frac{1}{2\tau^2}\boldsymbol{\beta}'\boldsymbol{\beta}\right\}I(h)$$

- If (5) is replaced in (14)

$$f(\boldsymbol{\beta}, \sigma^2/\boldsymbol{y}, k) = c(\sigma^2)^{-\frac{n+2}{2}}\left(\frac{\sigma^2}{k}\right)^{-\frac{p+2}{2}}$$

$$\times\, exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta}) - \frac{k}{2\sigma^2}\boldsymbol{\beta}'\boldsymbol{\beta}\right\} I(h)$$

$$= c(\sigma^2)^{-\frac{n+p+4}{2}} k^{\frac{p+2}{2}} \qquad (15)$$

$$\times\, exp\left\{-\frac{1}{2\sigma^2}\left[(\boldsymbol{y}-\boldsymbol{X\beta})'(\boldsymbol{y}-\boldsymbol{X\beta}) + k\boldsymbol{\beta}'\boldsymbol{\beta}\right]\right\} I(h)$$

- The parameter of interest in (15) is $\boldsymbol{\beta}$.
- When σ^2 is integrated out in $f(\boldsymbol{\beta}, \sigma^2/\boldsymbol{y}, k)$ (O'Donnell et al 2001):

$$f(\boldsymbol{\beta}/\boldsymbol{y}, k) = ck^{\frac{p+2}{2}}[(\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + k\boldsymbol{\beta}'\boldsymbol{\beta}]^{-\frac{n+p+2}{2}} I(h) \qquad (16)$$

- (15) and (16) assign positive probabilities on the parameter space only if the constraints are satisfied.
- Compare (6) and (16):

$$\mathcal{L}(\boldsymbol{\beta}; k) = (\boldsymbol{y} - \boldsymbol{X\beta})'(\boldsymbol{y} - \boldsymbol{X\beta}) + k\boldsymbol{\beta}'\boldsymbol{\beta} \qquad (6)$$

- The conditional expected value of β_j from (16) is

$$E(\beta_j/\boldsymbol{y},k) = \int\int\cdots\int \beta_j f(\boldsymbol{\beta}/\boldsymbol{y},k)d\beta_1 d\beta_2 \ldots d\beta_p \quad (17)$$

The estimate of this expected value is called the nonlinear inequality constrained ridge regression estimate of β_j.

- It is hard to evaluate (17) analytically.
- However, they can be numerically evaluated with MCMC.

MCMC estimate of $E(\beta_j/\boldsymbol{y},k)$:

- $\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, \ldots, \boldsymbol{\beta}^M$ sequence is produced from $f(\boldsymbol{\beta}/\boldsymbol{y},k)$.
- The definite integral in (17) is estimated with

$$\hat{\beta}_j(k) = \frac{1}{M}\sum_{t=1}^{M} \boldsymbol{\beta}^t \qquad (18)$$

  $\hat{\beta}_j(k)$ in (18) is the NICRR estimator of β_j.

- Algorithms such as Metropolis-Hastings, Gibbs sampling and adaptive harmonic mean can be used to produce the $\boldsymbol{\beta}^1, \boldsymbol{\beta}^2, \ldots, \boldsymbol{\beta}^M$ sequence (Casella and George, 1992; Chua et al, 2001).

The Metropolis-Hasting Algorithm:

1. Let $\boldsymbol{\beta}^0$ be an arbitrary starting point which satisfies $h_r(\boldsymbol{\beta}^0) \geq 0$ $(r = 1,2, \ldots, R)$. Set $i = 0$.
2. For the i_th $\boldsymbol{\beta}^i$, use the symmetrical transition density, $q\left(\boldsymbol{\beta}^i, \boldsymbol{\beta}^c\right)$, to compute the candidate parameter vector $\boldsymbol{\beta}^c$
3. If $h_r(\boldsymbol{\beta}^c) < 0$ for a restriction then set $\boldsymbol{\beta}^{i+1} = \boldsymbol{\beta}^i$ and go to Step 7.

The Metropolis-Hasting Algorithm (Continue):

4. Compute

$$\alpha(\boldsymbol{\beta}^i, \boldsymbol{\beta}^c) = min\left\{\frac{f(\boldsymbol{\beta}^c/\boldsymbol{y},k)}{f(\boldsymbol{\beta}^i/\boldsymbol{y},k)}, 1\right\}$$

with the marginal density function $f(\boldsymbol{\beta}/\boldsymbol{y}, k)$.

5. Generate a random number U from the continous uniform distribution between $[0,1]$.

6. Set $\boldsymbol{\beta}^{i+1} = \boldsymbol{\beta}^c$ if $U \leq \alpha(\boldsymbol{\beta}^i, \boldsymbol{\beta}^c)$, otherwise $\boldsymbol{\beta}^{i+1} = \boldsymbol{\beta}^i$.

7. Set $i = i + 1$ and go to Step 2.

The Metropolis-Hasting Algorithm (Continue):

- $N_p\left(\boldsymbol{\beta}^i, 0{,}2\hat{\sigma}^2\right)$ can be used for $q\left(\boldsymbol{\beta}^i, \boldsymbol{\beta}^c\right)$.
- k in (16) can be replaced with $\hat{k}_{HK} = p\hat{\sigma}^2/\widehat{\boldsymbol{\beta}}'\widehat{\boldsymbol{\beta}}$.
- $\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^s, \underbrace{\boldsymbol{\beta}^{s+1}, \ldots, \boldsymbol{\beta}^{s+M}}_{\substack{\text{Used}\\ \text{in the}\\ \text{estimation}}}$ will be generated.
- s: Burn-in period
- NICRR estimate of β_j:

$$\hat{\beta}_j(k) = \frac{1}{M}\sum_{t=s+1}^{s+M} \boldsymbol{\beta}^t$$

5. Monte Carlo Experiment

- OLS, RR, NICLS, and NICRR are compared.
- Scalar mean square error (mse) is used for the comparison.
- Independent variables are generated so that they are extent to multicollinearity.
- Different values for the parameters of the data generating process are used.

Model parameters:

- $n = 20, 40$
- $p = 5, 9$
- $\sigma^2 = 0{,}1; 1; 10$
- $R = 1, 2, 3$ ($p = 5$) and $1, 2, 3, 6$ ($p = 9$)
- For $p = 5$, $\beta = (2 \quad 3 \quad 5 \quad 7 \quad 10)'$
- For $p = 9$, $\beta = (2 \quad 3 \quad 5 \quad 7 \quad 10 \quad -2 \quad 6 \quad -4 \quad 8)'$

Restrictions:

- $R = 1$
  - $\beta_2\beta_3 - 14 \geq 0$
- $R = 2$
  - $\beta_2\beta_3 - 14 \geq 0$
  - $\beta_4/\beta_5 - 0{,}2 \geq 0$
- $R = 3$
  - $\beta_2\beta_3 - 14 \geq 0$
  - $\beta_4/\beta_5 - 0{,}2 \geq 0$
  - $\beta_4/\beta_2 - 1{,}5 \geq 0$
- $R = 6$
  - $\beta_2\beta_3 - 14 \geq 0$
  - $\beta_4/\beta_5 - 0{,}2 \geq 0$
  - $\beta_4/\beta_2 - 1{,}5 \geq 0$
  - $\beta_6^2\beta_1 - 7 \geq 0$
  - $\beta_8\beta_6 - 6 \geq 0$
  - $(\beta_9 + \beta_7)/\beta_5 - 0{,}5 \geq 0$

Independent variables:

- $\boldsymbol{x_1 = 1}$
- $z_{ij} \sim N(0,1),\ i = 1,2,\dots,n,\ j = 2,3,\dots,p+1$
- $x_{ij} = \sqrt{1-\rho^2} z_{ij} + \rho z_{i,p+1},\ i = 1,2,\dots,n,\ j = 2,3,\dots,p$ (Kibria, 2003)
- $corr(x_j, x_k) = \rho^2,\ (j \neq k)$
- $\rho = 0{,}8; 0{,}9; 0{,}99; 0{,}999; 0{,}9999$
- $\kappa = \sqrt{\lambda_{max}/\lambda_{min}}$

- $s = 300, M = 1000$
- $\boldsymbol{\beta}^0 = \boldsymbol{\beta}$

$$\widehat{hko}(\boldsymbol{\beta}^*) = \frac{1}{MCN}\sum_{mci=1}^{MCN}(\boldsymbol{\beta}^*_{mci} - \boldsymbol{\beta})'(\boldsymbol{\beta}^*_{mci} - \boldsymbol{\beta}) \qquad (19)$$

- $MCN = 1000$

- Simulations are done with MATLAB (The seed is = 2352367901235).

6. Findings

- There are 210 different possible combinations for different values of n, p, σ^2, R, ρ.
- The effect of each parameter is evaluated seperatly.

Tablo 1. The effect of ρ $(n = 40, p = 9, \sigma^2 = 10, R = 3)$

ρ	κ	OLS	RR	NICLS	NICRR
0,8	5,93	6,472	6,356	5,685	5,523
0,9	8,96	11,697	11,551	10,268	9,500
0,99	30,03	109,779	79,167	103,466	65,835
0,999	95,58	1108,173	374,716	316,630	158,884
0,9999	302,51	10568,763	2614,598	369,737	183,406

Tablo 2. The effect of R $(n = 20, p = 9, \sigma^2 = 1, \rho = 0{,}999)$

r	κ	OLS	RR	NICLS	NICRR
1	139,08	390,185	174,783	57,598	53,6142
2	139,08	400,063	178,602	53,398	54,2152
3	139,08	387,322	173,752	50,713	45,8364
4	139,08	386,773	177,756	46,754	28,4079

Tablo 3. The effect of σ^2 $(n = 20, p = 5, R = 3, \rho = 0{,}99)$

σ^2	κ	OLS	RR	NICLS	NICRR
0,1	21,92	1,347	1,319	1,142	1,194
1	21,92	13,629	11,399	9,655	8,526
10	21,92	134,730	60,071	56,148	32,355

Tablo 4. The effect of n and p ($\sigma^2 = 1, R = 3, \rho = 0{,}99$)

n	p	κ	OLS	RR	NICLS	NICRR
20	5	21,92	13,629	11,399	9,655	8,526
20	9	44,34	38,986	34,323	24,595	28,429
40	5	17,64	4,639	4,337	3,402	3,255
40	9	30,03	11,100	10,876	10,236	9,634

7. Results and Discussion

- mse of estimators are increasing with the degree of multicollinearity. NICRR is the least effected estimator from this situation.
- mse of estimators decrase as n increased, while increased as p increases.
- mse of estimators are increased with the error variance. NICRR is the least effected estimator from this situation.
- The performance of estimators are better as the number of constraints increase.
- Results show that NICRR generally outperms the remaining estimators and it has lower mse values.

Future research:

- The simulations will be supported with a real life application.
- The case where the constraints are not satisfied will be investigated.
- A likelihood-ratio test will be defined for the test of restrictions.
- An informative prior distribution (such as inverse chi-square, inverse gamma) for σ^2 and τ^2 will be used.
- Different starting points will be used instead of real $\boldsymbol{\beta}$.
- s, M and MCN will be increased.

References

- Casella, G., George, E.I. (1992), “Explaining the Gibbs sampler”, *The American Statistician*, 46, p167-173.
- Chua, C.L. , Griffiths, W.E. and O'Donnell, C.J. (2001), “Bayesian Model Averaging in Consumer Demand Systems with Inequality Constraints”, Australasian Meeting of The Econometric Society, July 6- 8, Auckland, New Zealand.
- Deaton, A. and Muellbauer, J. (1980), “An Almost Ideal Demand System”, *American Economic Review, 70*, p312-326.
- Geweke, J. (1986), “Exact Inference in the Inequality Constrained Normal Linear Regression Model”, *Journal of Applied Econometrics*, 1, p127-141.

References

- Geweke, J. (1999), “Using Simulation Methods For Bayesian Econometric Models: Inference, Development, and Communciation”, *Econometric Review*, 18, p1-73.
- Hastings, W.K. (1970), “Monte Carlo Sampling Methods using Markov Chains and their Applications”, *Biometrika, 57*, p97-109.
- Hoerl, A., and Kennard, R. (1970), Ridge regression: biased estimation for non-orthogonal problems. *Technometrics, 12*, 55-67.
- Jorgenson, D.W., Lau, L.J. and Stoker, T.M. (1982), “The Transcendental Logarithmic Model of Aggregate Consumer Behaviour”, *Advances in Economics (Vol 1),* eds. R.L Basmann and G. Rhodes, Greenwich, CT: JAI Press, p97-238.

References

- Kibria, B.M.G. (2003), "Performance of some new ridge regression estimators", *Communications in Statistics, Simulation and Computation,* 32(2), p419–435.
- Liew, C.K. (1976), "Inequality constrained least squares estimation", *Journal of the American Statistical Association*, 71(355), p746-751.
- Lindley, D.V. and Smith, A.F.M. (1972), "Bayes Estimates for the Linear Model", Journal of the Royal Statistical Society, Series B (Methodological), 34, p1-41.
- Metropolis, N., Rosenbluth, A.W. , Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953), "Equations of State Calculations by Fast Computing Machines", *Journal of Chemical Physics, 21*, p1087-1092.

References

- O'Donnell, C.J., Rambaldi, A.N. and Doran, H.E. (2001), "Estimating Economic Relationships Subject to Firm- and Time-Varying Equality and Inequality Constraints", *Journal of Applied Econometrics*, 17, p709-726.
- Toker, S., Üstündağ Şiray, G. and Kaçıranlar, S. (2013), "Inequality constrained ridge regression estimator ", *Statistics and Probability Letters*, 83, 2391-2398.

Thank you…