Firewall Authentication Keepalive ×   ยินดีต้อนรับผู้ใช้งาน   ×   Commonsense Explanations of S ×   How to Make a Decision Based ○ ×   +

link.springer.com/chapter/10.1007/978-3-030-77094-5_7?fbclid=IwAR2ttUKogStN3FdtvH4AKatr59Li93cyAuhEQZg4wL0XHDZDvkFZiw5CtCc

Apps   ซีรีย์เกาหลี Goblin ซับ...   THE STANDARD : S...   ซีรีย์เกาหลี One Sprin...   ซีรีย์เกาหลี Arthdal C...   360搜索热榜-你的...

🐦 **Springer** Link

Commonsense Explanations of Sparsity, Zipf Law, and Nash's Bargaining Solution

Authors	Authors and affiliations

Olga Kosheleva, Vladik Kreinovich ✉ , Kittawit Autchariyapanikul

Chapter
First Online: 27 July 2021

23
Downloads

Part of the Studies in Computational Intelligence book series (SCI, volume 983)

Abstract

As econometric models become more and more accurate and more and more mathematically complex, they also become less and less intuitively clear and convincing. To make these models more convincing, it is desirable to supplement the corresponding mathematics with commonsense explanations. In this paper, we provide such explanation for three economics-related concepts: sparsity (as in LASSO), Zipf's Law, and Nash's bargaining solution.

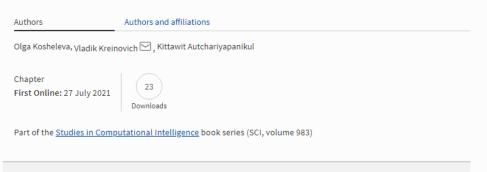# Commonsense Explanations of Sparsity, Zipf Law, and Nash's Bargaining Solution

International Econometric Conference of Vietnam

ECONVN 2021: Prediction and Causality in Econometrics and Related Topics pp 67-74 | Cite as

- Olga Kosheleva (1)
- Vladik Kreinovich (1) Email author (vladik@utep.edu)
- Kittawit Autchariyapanikul (2)

1. University of Texas at El Paso, , El Paso, USA
2. Maejo University, , Chiang Mai, Thailand

Chapter
First Online: 27 July 2021

- 23 Downloads

Part of the Studies in Computational Intelligence book series (SCI, volume 983)

Abstract

As econometric models become more and more accurate and more and more mathematically complex, they also become less and less intuitively clear and convincing. To make these models more convincing, it is desirable to supplement the corresponding mathematics with commonsense explanations. In this paper, we provide such explanation for three economics-related concepts: sparsity (as in LASSO), Zipf's Law, and Nash's bargaining solution.

This is a preview of subscription content, log in to check access.

Notes

Acknowledgments

References

1. Alkhatib, H., Neumann, I., Kreinovich, V., Van Le, C.: Why LASSO, EN, and CLOT: invariance-based explanation. In: Trung, N.D., Thach, N.N., Kreinovich, V. (eds.) Data Science for Financial Econometrics, pp. 37–50. Springer (2021). https://doi.org/10.1007/978-3-030-48853-6_2 (https://doi.org/10.1007/978-3-030-48853-6_2)

2. Bellamy, R.K.E., et al.: Think your Artifical Intelligence software is fair? Thing again, Computing Edge, pp. 14–18, February 2020
Google Scholar (https://scholar.google.com/scholar?q=Bellamy%2C%20R.K.E.%2C%20et%20al.%3A%20Think%20your%20Artifical%20Intelligence%20software%20is%20fair%3F%20Thing%20again%2C%20Computing%20Edge%2C%20pp.%2014%E2%80%9318%2C%20February%202020)

3. Belohlavek, R., Dauben, J.W., Klir, G.J.: Fuzzy Logic and Mathematics: A Historical Perspective. Oxford University Press, New York (2017)
CrossRef (https://doi.org/10.1093/oso/9780190200015.001.0001)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Fuzzy%20Logic%20and%20Mathematics%3A%20A%20Historical%20Perspective&author=R.%20Belohlavek&author=JW.%20Dauben&author=GJ.%20Klir&publication_year=2017)

4. Amizic, B., Spinoulas, L., Molina, R., Katsaggelos, A.K.: Compressive blind image deconvolution. IEEE Trans. Image Process. **22**(10), 3994–4006 (2013)
MathSciNet (http://www.ams.org/mathscinet-getitem?mr=3105158)
CrossRef (https://doi.org/10.1109/TIP.2013.2266100)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Compressive%20blind%20image%20deconvolution&author=B.%20Amizic&author=L.%20Spinoulas&author=R.%20Molina&author=AK.%20Katsaggelos&journal=IEEE%20Trans.%20Image%20Process.&volume=22&issue=10&pages=3994-4006&publication_year=2013)

5. Candès, E.J., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Comm. Pure Appl. Math. **59**, 1207–1223 (2006)
MathSciNet (http://www.ams.org/mathscinet-getitem?mr=2230846)
CrossRef (https://doi.org/10.1002/cpa.20124)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Stable%20signal%20recovery%20from%20incomplete%20and%20inaccurate%20measurements&author=EJ.%20Cand%C3%A8s&author=J.%20Romberg&author=T.%20Tao&journal=Comm.%20Pure%20Appl.%20Math.&volume=59&pages=1207-1223&publication_year=2006)

6. Candès, E., Romberg, J., Tao, T.: Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theor. **52**(2), 489–509 (2006)
MathSciNet (http://www.ams.org/mathscinet-getitem?mr=2236170)
CrossRef (https://doi.org/10.1109/TIT.2005.862083)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Robust%20uncertainty%20principles%3A%20Exact%20signal%20reconstruction%20from%20highly%20incomplete%20frequency%20information&author=E.%20Cand%C3%A8s&author=J.%20Romberg&author=T.%20Tao&journal=IEEE%20Trans.%20Inf.%20Theor.&volume=52&issue=2&pages=489-509&publication_year=2006)

7. Candès, E.J., Tao, T.: Decoding by linear programming. IEEE Trans. Inf. Theor. **51**(12), 4203–4215 (2005)
MathSciNet (http://www.ams.org/mathscinet-getitem?mr=2243152)

CrossRef (https://doi.org/10.1109/TIT.2005.858979)
Google Scholar (http://scholar.google.com/scholar_lookup?
title=Decoding%20by%20linear%20programming&author=EJ.%20Cand%C3%
A8s&author=T.%20Tao&journal=IEEE%20Trans.%20Inf.%20Theor.&volume
=51&issue=12&pages=4203-4215&publication_year=2005)

8. Candès, E.J., Wakin, M.B.: An Introduction to compressive sampling. IEEE
Signal Process. Mag. **25**(2), 21–30 (2008)
CrossRef (https://doi.org/10.1109/MSP.2007.914731)
Google Scholar (http://scholar.google.com/scholar_lookup?
title=An%20Introduction%20to%20compressive%20sampling&author=EJ.%2
0Cand%C3%A8s&author=MB.%20Wakin&journal=IEEE%20Signal%20Proces
s.%20Mag.&volume=25&issue=2&pages=21-30&publication_year=2008)

9. Cervantes, D., Kosheleva, O., Kreinovich, V.: Why Zipf's law: a symmetry-based
explanation. Int. Math. Forum **13**(6), 255–258 (2018)
CrossRef (https://doi.org/10.12988/imf.2018.8316)
Google Scholar (http://scholar.google.com/scholar_lookup?
title=Why%20Zipf%E2%80%99s%20law%3A%20a%20symmetry-
based%20explanation&author=D.%20Cervantes&author=O.%20Kosheleva&au
thor=V.%20Kreinovich&journal=Int.%20Math.%20Forum&volume=13&issue=
6&pages=255-258&publication_year=2018)

10. Cervantes, F., Usevitch, B., Valera, L., Kreinovich, V.: Why sparse? Fuzzy
techniques explain empirical efficiency of sparsity-based data- and image-
processing algorithms. In: Proceedings of the 2016 World Conference on Soft
Computing, Berkeley, California, pp. 165–169, 22–25 May 2016
Google Scholar (https://scholar.google.com/scholar?
q=Cervantes%2C%20F.%2C%20Usevitch%2C%20B.%2C%20Valera%2C%20L.
%2C%20Kreinovich%2C%20V.%3A%20Why%20sparse%3F%20Fuzzy%20tech
niques%20explain%20empirical%20efficiency%20of%20sparsity-
based%20data-%20and%20image-
processing%20algorithms.%20In%3A%20Proceedings%20of%20the%202016
%20World%20Conference%20on%20Soft%20Computing%2C%20Berkeley%2
C%20California%2C%20pp.%20165%E2%80%93169%2C%2022%E2%80%932
5%20May%202016)

11. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theor. **52**(4), 1289–1306
(2005)
MathSciNet (http://www.ams.org/mathscinet-getitem?mr=2241189)
CrossRef (https://doi.org/10.1109/TIT.2006.871582)
Google Scholar (http://scholar.google.com/scholar_lookup?
title=Compressed%20sensing&author=DL.%20Donoho&journal=IEEE%20Tra
ns.%20Inf.%20Theor.&volume=52&issue=4&pages=1289-
1306&publication_year=2005)

12. Duarte, M.F., et al.: Single-pixel imaging via compressive sampling. IEEE
Signal Process. Mag. **25**(2), 83–91 (2008)
MathSciNet (http://www.ams.org/mathscinet-getitem?mr=2440485)
CrossRef (https://doi.org/10.1109/MSP.2007.914730)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Single-
pixel%20imaging%20via%20compressive%20sampling&author=MF.%20Duart
e&journal=IEEE%20Signal%20Process.%20Mag.&volume=25&issue=2&pages
=83-91&publication_year=2008)

13. Dumrongpokaphan, T., Kosheleva, O., Kreinovich, V., Belina, A.: Why sparse?
In: Kosheleva, O., Shary, S., Xiang, G., Zapatrin, R. (eds.) Beyond Traditional

Probabilistic Data Processing Techniques: Interval, Fuzzy, etc, pp. 461–470. Methods and Their Applications, Springer, Cham, Switzerland (2020). https://doi.org/10.1007/978-3-030-31041-7 (https://doi.org/10.1007/978-3-030-31041-7)

14. Edeler, T., Ohliger, K., Hussmann, S., Mertins, A.: Super-resolution model for a compressed-sensing measurement setup. IEEE Trans. Instrum. Meas. **61**(5), 1140–1148 (2012)
CrossRef (https://doi.org/10.1109/TIM.2011.2174891)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Super-resolution%20model%20for%20a%20compressed-sensing%20measurement%20setup&author=T.%20Edeler&author=K.%20Ohliger&author=S.%20Hussmann&author=A.%20Mertins&journal=IEEE%20Trans.%20Instrum.%20Meas.&volume=61&issue=5&pages=1140-1148&publication_year=2012)

15. Elad, M.: Sparse and Redundant Representations. Springer Verlag (2010). https://doi.org/10.1007/978-1-4419-7011-4_9 (https://doi.org/10.1007/978-1-4419-7011-4_9)

16. Hastie, T., Tibshirani, R., Wainwright, M.: Statistical Learning with Sparsity: The Lasso and Generalizations. Chapman and Hall/CRC, Boca Raton (2015)
CrossRef (https://doi.org/10.1201/b18401)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Statistical%20Learning%20with%20Sparsity%3A%20The%20Lasso%20and%20Generalizations&author=T.%20Hastie&author=R.%20Tibshirani&author=M.%20Wainwright&publication_year=2015)

17. Jaynes, E.T., Bretthorst, G.L.: Probability Theory: The Logic of Science. Cambridge University Press, Cambridge (2003)
CrossRef (https://doi.org/10.1017/CBO9780511790423)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Probability%20Theory%3A%20The%20Logic%20of%20Science&author=ET.%20Jaynes&author=GL.%20Bretthorst&publication_year=2003)

18. Klir, G., Yuan, B.: Fuzzy Sets and Fuzzy Logic. Prentice Hall, Upper Saddle River (1995)
zbMATH (http://www.emis.de/MATH-item?0915.03001)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Fuzzy%20Sets%20and%20Fuzzy%20Logic&author=G.%20Klir&author=B.%20Yuan&publication_year=1995)

19. Luce, R.D., Raiffa, R.: Games and Decisions: Introduction and Critical Survey. Dover, New York (1989)
zbMATH (http://www.emis.de/MATH-item?1233.91002)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Games%20and%20Decisions%3A%20Introduction%20and%20Critical%20Survey&author=RD.%20Luce&author=R.%20Raiffa&publication_year=1989)

20. Ma, J., Le Dimet, F.-X.: Deblurring from highly incomplete measurements for remote sensing. IEEE Trans. Geosci. Remote Sens. **47**(3), 792–802 (2009)
CrossRef (https://doi.org/10.1109/TGRS.2008.2004709)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Deblurring%20from%20highly%20incomplete%20measurements%20for%20remote%20sensing&author=J.%20Ma&author=F-X.%20Dimet&journal=IEEE%20Trans.%20Geosci.%20Remote%20Sens.&volume=47&issue=3&pages=792-802&publication_year=2009)

21. Mandelbrot, B.: The Fractal Geometry of Nature. Freeman, San Francisco, California (1983)
    Google Scholar (https://scholar.google.com/scholar?
    q=Mandelbrot%2C%20B.%3A%20The%20Fractal%20Geometry%20of%20Nat
    ure.%20Freeman%2C%20San%20Francisco%2C%20California%20%281983%
    29)

22. McMackin, L., Herman, M.A., Chatterjee, B., Weldon, M.: A high-resolution SWIR camera via compressed sensing. Proc. SPIE **8353**(1), 835303 (2012)
    CrossRef (https://doi.org/10.1117/12.920050)
    Google Scholar (http://scholar.google.com/scholar_lookup?title=A%20high-
    resolution%20SWIR%20camera%20via%20compressed%20sensing&author=L
    .%20McMackin&author=MA.%20Herman&author=B.%20Chatterjee&author=
    M.%20Weldon&journal=Proc.%20SPIE&volume=8353&issue=1&pages=83530
    3&publication_year=2012)

23. Mendel, J.M.: Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions. Springer, Cham, Switzerland (2017). https://doi.org/10.1007/978-
    3-319-51370-6 (https://doi.org/10.1007/978-3-319-51370-6)

24. Mittelstadt, B., Russell, C., Wachter, S.: Explaining explanations in AI. In: Proceedings of the 2019 ACM Fairness, Accounatbility, and Transparency Conference FAT 2019, Atlanta, Georgia, pp. 279–288, 29–31 January 2019
    Google Scholar (https://scholar.google.com/scholar?
    q=Mittelstadt%2C%20B.%2C%20Russell%2C%20C.%2C%20Wachter%2C%20
    S.%3A%20Explaining%20explanations%20in%20AI.%20In%3A%20Proceedin
    gs%20of%20the%202019%20ACM%20Fairness%2C%20Accounatbility%2C%2
    0and%20Transparency%20Conference%20FAT%202019%2C%20Atlanta%2C
    %20Georgia%2C%20pp.%20279%E2%80%93288%2C%2029%E2%80%9331%
    20January%202019)

25. Nash, J.: The bargaining problem. Econometrica **18**(2), 155–162 (1950)
    MathSciNet (http://www.ams.org/mathscinet-getitem?mr=35977)
    CrossRef (https://doi.org/10.2307/1907266)
    Google Scholar (http://scholar.google.com/scholar_lookup?
    title=The%20bargaining%20problem&author=J.%20Nash&journal=Economet
    rica&volume=18&issue=2&pages=155-162&publication_year=1950)

26. Natarajan, B.K.: Sparse approximate solutions to linear systems. SIAM J. Comput. **24**, 227–234 (1995)
    MathSciNet (http://www.ams.org/mathscinet-getitem?mr=1320206)
    CrossRef (https://doi.org/10.1137/S0097539792240406)
    Google Scholar (http://scholar.google.com/scholar_lookup?
    title=Sparse%20approximate%20solutions%20to%20linear%20systems&autho
    r=BK.%20Natarajan&journal=SIAM%20J.%20Comput.&volume=24&pages=2
    27-234&publication_year=1995)

27. Nguyen, H.P., Bokati, L., Kreinovich, V.: New (simplified) derivation of Nash's bargaining solution. J. Adv. Comput. Intell. Intell. Inform. (JACIII) **24**(5), 589–592 (2020)
    Google Scholar (https://scholar.google.com/scholar?
    q=Nguyen%2C%20H.P.%2C%20Bokati%2C%20L.%2C%20Kreinovich%2C%20
    V.%3A%20New%20%28simplified%29%20derivation%20of%20Nash%E2%80
    %99s%20bargaining%20solution.%20J.%20Adv.%20Comput.%20Intell.%20In
    tell.%20Inform.%20%28JACIII%29%2024%285%29%2C%20589%E2%80%93
    592%20%282020%29)

28. Nguyen, H.T., Walker, C.L., Walker, E.A.: A First Course in Fuzzy Logic. Chapman and Hall/CRC, Boca Raton (2019)
zbMATH (http://www.emis.de/MATH-item?1405.03001)
Google Scholar (http://scholar.google.com/scholar_lookup?title=A%20First%20Course%20in%20Fuzzy%20Logic&author=HT.%20Nguyen&author=CL.%20Walker&author=EA.%20Walker&publication_year=2019)

29. Novák, V., Perfilieva, I., Močkoř, J.: Mathematical Principles of Fuzzy Logic. Kluwer, Boston, Dordrecht (1999)
CrossRef (https://doi.org/10.1007/978-1-4615-5217-8)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Mathematical%20Principles%20of%20Fuzzy%20Logic&author=V.%20Nov%C3%A1k&author=I.%20Perfilieva&author=J.%20Mo%C4%8Dko%C5%99&publication_year=1999)

30. O'Neil, C.: Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Crown Books, New York (2016)
zbMATH (http://www.emis.de/MATH-item?1441.00001)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Weapons%20of%20Math%20Destruction%3A%20How%20Big%20Data%20Increases%20Inequality%20and%20Threatens%20Democracy&author=C.%20O%E2%80%99Neil&publication_year=2016)

31. Tibshirani, R.: Regression shrinkage and selection via the LASSO. J. Royal Stat. Soc. **58**(1), 267–288 (1996)
MathSciNet (http://www.ams.org/mathscinet-getitem?mr=1379242)
zbMATH (http://www.emis.de/MATH-item?0850.62538)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Regression%20shrinkage%20and%20selection%20via%20the%20LASSO&author=R.%20Tibshirani&journal=J.%20Royal%20Stat.%20Soc.&volume=58&issue=1&pages=267-288&publication_year=1996)

32. Tsaig, Y., Donoho, D.: Compressed sensing. IEEE Trans. Inf. Theor. **52**(4), 1289–1306 (2006)
MathSciNet (http://www.ams.org/mathscinet-getitem?mr=2241189)
CrossRef (https://doi.org/10.1109/TIT.2006.871582)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Compressed%20sensing&author=Y.%20Tsaig&author=D.%20Donoho&journal=IEEE%20Trans.%20Inf.%20Theor.&volume=52&issue=4&pages=1289-1306&publication_year=2006)

33. Xiao, L., Shao, J., Huang, L., Wei, Z.: Compounded regularization and fast algorithm for compressive sensing deconvolution. In: Proceedings of the 6th International Conference on Image Graphics, pp. 616–621 (2011)
Google Scholar (https://scholar.google.com/scholar?q=Xiao%2C%20L.%2C%20Shao%2C%20J.%2C%20Huang%2C%20L.%2C%20Wei%2C%20Z.%3A%20Compounded%20regularization%20and%20fast%20algorithm%20for%20compressive%20sensing%20deconvolution.%20In%3A%20Proceedings%20of%20the%206th%20International%20Conference%20on%20Image%20Graphics%2C%20pp.%20616%E2%80%93621%20%282011%29)

34. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)
CrossRef (https://doi.org/10.1016/S0019-9958(65)90241-X)
Google Scholar (http://scholar.google.com/scholar_lookup?title=Fuzzy%20sets&author=LA.%20Zadeh&journal=Inf.%20Control&volume=8&pages=338-353&publication_year=1965)

Copyright information

About this chapter

Personalised recommendations

Commonsense Explanations of Sparsity, Zipf Law, and Nash's Bargaining Solution

Olga Kosheleva

Vladik Kreinovich

Kittawit Autchariyapanitkul

Commonsense Explanations of Sparsity, Zipf Law, and Nash's Bargaining Solution

Olga Kosheleva[1], Vladik Kreinovich[1], and
Kittawit Autchariyapanikul[3]
[1]University of Texas at El Paso
500 W. University
El Paso, YX 79968, USA
olgak@utep.edu. vladik@utep.edu
[2]Maejo University, Chiang Mai, Thailand
kittawit.a@mju.ac.th

Abstract

As econometric models become more and more accurate and more and more mathematically complex, they also become less and less intuitively clear and convincing. To make these models more convincing, it is desirable to supplement the corresponding mathematics with commonsense explanations. In this paper, we provide such explanation for three economics-related concepts: sparsity (as in LASSO), Zipf's Law, and Nash's bargaining solution.

1 Formulation of the Problem

Need for commonsense explanations. In the 19 century, from Adam Smith through Karl Marx, economic phenomena were described by relatively simple intuitively clear formulas. However, it was well understood that these formulas provide, at best, a qualitative understanding, and qualitative predictions.

Economics and finance are complex phenomena, so not surprisingly, simple models do not provide a good quantitative description of these phenomena. Starting with the 20 century, new models appeared that allow us, in many cases, to make reasonable accurate predictions. The corresponding models – that more adequately describe economic and financial phenomena – are often very mathematically complicated, far from common sense.

And here lies a problem. Since models are complicated, requiring complex math beyond what most people know, people in general – and politicians in particular – do not fully trust these models. This lack of complete trust is understandable, since sometimes, predictions of previous models turn out to be wrong, and decisions based on these predictions made crisis situations worse. It

is therefore desirable to come up with commonsense explanations for the current models – or at least for the major assumptions behind these models and for the main features of these models.

This need is similar to the well-publicized need for explainable AI (see, e.g., [2, 24, 30]) – for example, a female entrepreneur who was denied a bank loan by an AI-computer program would like to make sure that this denial was based on inadequate state of her company and not on the fact that, since in the past, most loans were given to men, the statistics-based system automatically continues this trend.

What we do in this paper. Of course, the goal of having explainable econometrics is complex, it requires joint efforts of many researchers. In this paper, we provide just three examples of the desired commonsense explanations: namely, we provide commonsense explanation for sparsity, for Zipf law, and for Nash's bargaining solution.

2 Commonsense Explanation of Sparsity

What is sparsity: a brief reminder. Econometric analysis often starts with trying to use linear regression: we try to predict the value $y(t+1)$ of the desirable quantity y at the next moment of time $t + 1$ (e.g., next month or next year) as a linear combination of the values $x_1(t)$, ..., $x_n(t)$ of different quantities at the current moment of time – and maybe of their past values $x_i(t - 1)$, $x_i(t - 2)$, ...

Usually, just to be on the safe side, we add as many quantities x_i and as many past moments of time as possible, realizing that some of these values may be irrelevant. In the process of regression, it usually turns out that many of these variables are indeed irrelevant – we can see it by observing that the coefficients at these variables are close to 0.

The problem with using the usual linear regression methods for linear regression is that, since the data is noisy, and linear dependence is approximate, we do not get perfect match, and thus, do not get exactly 0 values even where there is no dependence, and these erroneously non-zero value make the resulting predictions much less accurate. Much more accurate predictions can be obtained if we explicitly require, from the very beginning, that a large number of the coefficients be zero, i.e., that the dependence is *sparse*. This is the main idea behind, e.g., the LASSO method [1, 16, 31].

Comment. Sparsity is not limited to economics and finance, it is a useful tool in data processing in general; see, e.g., [4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 20, 22, 26, 32, 33].

Why do we need a commonsense explanation for sparsity. At first glance, what we described makes sense, but one can also raise a serious counter-argument: in economics and in finance, everything is inter-related, often, a small event across the world causes some stocks to go up or down, so why would there be zero coefficients at all?

From this viewpoint, if one looks seriously at the previous text, this text does not provide any commonsense explanation, just states the empirical fact that the resulting best-fit formulas are often sparse.

It would thus be nice to have some commonsense explanation for this empirical phenomenon of sparsity.

Continuity: main commonsense feature underlying our explanation. Our commonsense explanation is based on a reasonable feature of *continuity* – that small changes in one or several quantities should lead to small changes in other quantities.

Of course, there are extreme situations when there is no continuity: e.g., if the amount of money that a company needs to pay right away is slightly smaller than its current amount of liquid assets, it can easily pay its creditors and suppliers, while if the needed amount is slightly larger than what is available, the company may be in trouble. But even this situation is more purely theoretical: if the difference is very small, the company can easily take a loan or negotiate a payment extension with its suppliers – and if they refuse because of the bad financial history of this company, this means that the trouble is in this bad history and *not* in the minor inability to fully pay the debts right away.

Let us describe continuity in precise terms. Suppose that we have a linear relation between the input quantities $x_1, \ldots, x_n$ and the desired quantity y:

$$y = a_0 + a_1 \cdot x_1 + \ldots + a_n \cdot x_n. \tag{1}$$

Continuity means, in particular, that if we make a small change in one or several values x_i, i.e., replace the original value x_i with a new value $x_i + \Delta x_i$, where $|\Delta x_i| \leq \varepsilon$ for some $\varepsilon > 0$, then the resulting change Δy in y must also be small, i.e., smaller than $\delta > 0$ for some $\delta \approx \varepsilon$. (It may be tempting to take $\delta = \varepsilon$, but, as we will see, this is not possible.)

Similarly, if we analyze the dependence of y on one of the variables x_i – i.e., if we analyze how much we need to change x_i to reach the desired change in y – then we can formulate a similar requirement: that for each small change Δy (when $|\Delta y| \leq \varepsilon$), the corresponding change Δx_i should also be small: $|\Delta x_i| \leq \delta$.

Comment. Of course, this requirement only makes sense only when a change in x_i can actually change y, i.e., when the corresponding coefficient a_i is different from 0.

What we can conclude from continuity. If we change one of the variables x_i by $\Delta x_i = \varepsilon$, then the value of y changes by $\Delta y = a_i \cdot \Delta x_i = a_i \cdot \varepsilon$. So, the requirement that $|\Delta y| \leq \delta$ takes the form $|a_i \cdot \varepsilon| \leq \delta$, i.e., that

$$|a_i| \leq \frac{\delta}{\varepsilon}. \tag{2}$$

Similarly, for the indices i for which $a_i \neq 0$, the change in y caused by a change Δx_i in x_i is equal to $\Delta y = a_i \cdot \Delta x_i$. Thus, to get the change $\Delta y = \varepsilon$,

we need to take $\Delta x_i = \dfrac{\Delta y}{a_i} = \dfrac{\varepsilon}{a_i}$. The requirement that $|\Delta x_i| \le \delta$ thus takes the form $\left| \dfrac{\varepsilon}{a_i} \right| \le \delta$, i.e., equivalently,

$$\frac{\varepsilon}{\delta} \le |a_i|. \tag{3}$$

This, by the way, shows why we cannot take $\delta = \varepsilon$ – in this case, from (2) and (3), we would conclude that $|a_i| \le 1$ and $1 \le |a_i|$, so $|a_i| = 1$ – but empirically, we can have different values a_i. Moreover, we must have $\varepsilon < \delta$ – otherwise, if $\varepsilon > \delta$, then (2) and (3) would lead to two inconsistent inequalities $|a_i| < 1$ and $1 < |a_i|$.

So far, we considered the case when we change only one input. What if we change all of them, e.g., if we take $\Delta x_i = \mathrm{sign}(a_i) \cdot \varepsilon$ for all i? In this case, we have

$$\Delta y = \sum_{i=1}^{n} a_i \cdot \mathrm{sign}(a_i) \cdot \varepsilon = \sum_{i=1}^{n} |a_i| \cdot \varepsilon = \varepsilon \cdot \sum_{i=1}^{n} |a_i|.$$

In this formula, we can safely ignore all the indices for which $a_i = 0$, so

$$\Delta y = \varepsilon \cdot \sum_{i:a_i \ne 0} |a_i|.$$

Our continuity requirement is that $|\Delta y| \le \delta$. In this example, $\Delta y > 0$, so $|\Delta y| = \Delta y$, and the desired inequality takes the form

$$\varepsilon \cdot \sum_{i:a_i \ne 0} |a_i| \le \delta. \tag{4}$$

We already know that $\dfrac{\varepsilon}{\delta} \le |a_i|$, thus,

$$n_i \cdot \frac{\varepsilon}{\delta} \le \sum_{i:a_i \ne 0} |a_i|, \tag{5}$$

where by n_i, we denotes the number of non-zero coefficients a_i. Multiplying both side of inequality (5) by ε, we conclude that

$$n_i \cdot \frac{\varepsilon^2}{\delta} \le \varepsilon \cdot \sum_{i:a_i \ne 0} |a_i|. \tag{6}$$

By combining inequalities (4) and (6), we conclude that

$$n_i \cdot \frac{\varepsilon^2}{\delta} \le \delta,$$

i.e., equivalently, that

$$n_i \le \frac{\delta^2}{\varepsilon^2}.$$

Since $\delta \approx \varepsilon$, this means that the ratio $\dfrac{\delta^2}{\varepsilon^2}$ is reasonably small, so the number n_i of non-zero coefficients is small – which is exactly sparsity!

3   Commonsense Explanation of Zipf's Law

Zipf's law: a brief reminder. This law is named after a scientist who discovered that in linguistics: if we order all the words from a natural language from the most frequent to the least frequent ones, then the frequency p_n of the n-th word in this order is approximately equal to c/n. This law turned out to be ubiquitous, it is applicable to many phenomena, including economic ones; see, e.g., [9, 21]. For example, it describes the distribution of companies by size.

How can we explain this law in commonsense terms?

Main idea behind our explanation. The fact that we sorted objects (words, companies, etc.) in the reverse order according to their frequency means that:

- if $m < n$, then $p_n < p_m$, and

- vice versa, if $p_n < p_m$, then $m < n$.

These formal statements are not very informative: they are satisfied for Zipf's law $p_n = c/n$, but they could be satisfied for all other possible monotonic sequences, e.g., $p_n = c/n^2$ or $p_n = \exp(-c \cdot n)$.

However, a natural commonsense understanding of these two conditions goes beyond their formal definition. It is natural, e.g., to also conclude that m is somewhat smaller than n, then p_n is somewhat smaller than p_m, and that if m is much smaller than n, then p_n is much smaller than p_m.

This commonsense understanding can be described as follows: intuitively, a statement that A implies B means not only that if A is true then B is true, it also means that, more generally, the degree $d(B)$ to which we believe in B is at least as large as the degree $d(A)$ to which we believe in A.

Comment. In principle, we could formalize the notion of a degree – e.g., by using fuzzy logic [3, 18, 23, 28, 29, 34] – but let us stick to commonsense ideas, and do now go into formalization unless it becomes necessary.

Resulting explanation. From the viewpoint of the above understanding, the two implication statements mean the following:

- the first statement means that the degree $d(p_n < p_m)$ to which we believe that $p_n < p_m$ is greater than or equal to the degree of belief $d(m < n)$ that $m < n$:
$$d(m < n) \leq d(p_n < p_m);$$

- similarly, the second statement means that the degree $d(m < n)$ to which we believe that $m < n$ is greater than or equal to the degree of belief $d(p_n < p_m)$ that $p_n < p_m$:
$$d(p_n < p_m) \leq d(m < n).$$

The above two inequalities imply that these degree are equal:
$$d(p_n < p_m) = d(m < n). \tag{7}$$

How do we describe this degree? In principle, this can be done differently, but in economic applications, this is usually understood in relative – percent – terms. If we say that someone started earning 10K more per year, we are not explaining anything: it may be going to 0 to 10K, or it may be a miserable increase from 300K to 310K. A much better commonsense understanding is provided by the ratio of two numbers: no matter what we started with, a 3% increase is not much, a 20% increase is always significant, and a 100% increase is always drastic.

Thus, in economic applications, the degree $d(a < b)$ to which a is smaller than b is determined only by the ratio, i.e., $d(a < b) = f(b/a)$ for some increasing function $f(x)$. From this viewpoint, the formula (7) takes the form

$$f\left(\frac{p_m}{p_n}\right) = f\left(\frac{n}{m}\right).$$

Since the function $f(x)$ is increasing – the larger the ratio b/a, the more we are convincing that b is larger – this equality is equivalent to

$$\frac{p_m}{p_n} = \frac{n}{m}.$$

If we move all the terms containing m to one side and all the terms containing n to the other side, we conclude that $m \cdot p_m = n \cdot p_n$ for all m and n. This product is the same for all n, i.e., is a constant: $n \cdot p_n = c$. Thus, $p_n = c/n$, exactly the Zipf's law!

4 Commonsense Explanation of Nash's Bargaining Solution

Nash's bargaining solution: a brief reminder. In many real-life situations, the outcome depends on the actions of several actors. A classical example is several countries producing oil. Depending on their actions, they may get different gains $(g_1, \ldots, g_n)$. In some possible outcomes, one actor gets more, in other outcomes, another actor gets more. For example, if one country cuts off its production and others don't, other countries will benefit.

If the actors do not coordinate their action – e.g., if we all produce too much oil, the price of oil will dive, and they will all lose. So, they need to come up with a joint compromise solution.

A Nobelist John Nash used some – rather complex – math – to show that the best strategy is to maximize the product $\prod_{i=1}^{n} g_i$ of the participants' gains; see, e.g., [19, 25, 27].

It is desirable to have a commonsense explanation. From the mathematical viewpoint, it is a good solution, but from the commonsense viewpoint, multiplying gains does not make any sense.

How can we explain this idea in commonsense terms?

Main idea and the resulting explanation. How can we gauge the quality of a solution? For athletes, a natural measure of the athlete's quality is the number of other athletes whose performance is lower. For standardized tests, the person's results are usually described by the percentage of test takers who had worse performance. This way it is clear that if this percentage is 3%, the graduate program applicant is not very good, while if it is 99.9%, we should accept this person right away.

Similarly, we can gauge the quality of an alternative $g = (g_1, \ldots, g_n)$ by counting how many alternatives are worse, i.e., how many there are tuples $g' = (g'_1, \ldots, g'_n) \neq g$ which are clearly worse, i.e., for which $g'_i \leq g_i$ for all i (since $g' \neq g$, this implies that $g'_i < g_i$ for some i).

From the purely mathematical viewpoint, there are infinitely many such tuples. In practice, however, there is some "quantum" here, the smallest amount $\varepsilon > 0$ below which the difference is not noticeable. It may be 1 cent if we talk about people, it may be 1 million dollars if we talk about oil producing companies.

From this viewpoint, there are g_1/ε values 0, ε, 2ε, ... which are smaller than or equal to g_1. For each of these values, there are g_2/ε values which are smaller than or equal to g_2 – i.e., $(g_1/\varepsilon) \cdot (g_2/\varepsilon)$ possible combinations. Similarly, we conclude that overall, there are

$$(g_1/\varepsilon) \cdot \ldots \cdot (g_n/\varepsilon) = \frac{\prod_{i=1}^{n} g_i}{\varepsilon^n}$$

possible alternatives which are worse than the alternative $g = (g_1, \ldots, g_n)$. Maximizing this number is equivalent to maximizing the product $\prod_{i=1}^{n} g_i$ – which is exactly Nash's bargaining solution!

Acknowledgments

References

[1] H. Alkhatib, I. Neumann, V. Kreinovich, and Chon Van Le, "Why LASSO, EN, and CLOT: invariance-based explanation", In: Nguyen Duc Trung, Nguyen Ngoc Thach, and V. Kreinovich (eds.), *Data Science for Financial Econometrics*, Springer, to appear.

[2] R. K. E. Bellamy et al., "Think your Artifical Intelligence software is fair? Thing again", *Computing Edge*, February 2020, pp. 14–18.

[3] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[4] B. Amizic, L. Spinoulas, R. Molina, and A. K. Katsaggelos, "Compressive blind image deconvolution", *IEEE Transactions on Image Processing*, 2013, Vol. 22, No. 10, pp. 3994–4006.

[5] E. J. Candès, J. Romberg and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements", *Comm. Pure Appl. Math.*, 2006, Vol. 59, pp. 1207–1223.

[6] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information", *IEEE Transactions on Information Theory*, 2006, Vol. 52, No. 2, pp. 489–509.

[7] E. J. Candès and T. Tao, "Decoding by linear programming", *IEEE Transactions on Information Theory*, 2005, Vol. 51, No. 12, pp. 4203–4215.

[8] E. J. Candès and M. B. Wakin, "An Introduction to compressive sampling", *IEEE Signal Processing Magazine*, 2008, Vol. 25, No. 2, pp. 21–30.

[9] D. Cervantes, O. Kosheleva, and V. Kreinovich, "Why Zipf's law: a symmetry-based explanation", *International Mathematical Forum*, 2018, Vol. 13, No. 6, pp. 255–258.

[10] F. Cervantes, B. Usevitch, L. Valera, and V. Kreinovich, "Why sparse? fuzzy techniques explain empirical efficiency of sparsity-based data- and image-processing algorithms", *Proceedings of the 2016 World Conference on Soft Computing*, Berkeley, California, May 22–25, 2016, pp. 165–169.

[11] D. L. Donoho, "Compressed sensing", *IEEE Transactions on Information Theory*, 2005, Vol. 52, No. 4, pp. 1289–1306.

[12] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, 2008, Vol. 25, No. 2, pp. 83–91.

[13] T. Dumrongpokaphan, O. Kosheleva, V. Kreinovich, and A. Belina, "Why sparse?", In: O. Kosheleva, S. Shary, G. Xiang, and R. Zapatrin (eds.), *Beyond Traditional Probabilistic Data Processing Techniques: Interval, Fuzzy, etc. Methods and Their Applications*, Springer, Cham, Switzerland, 2020, pp. 461–470.

[14] T. Edeler, K. Ohliger, S. Hussmann, and A. Mertins, "Super-resolution model for a compressed-sensing measurement setup", *IEEE Transactions on Instrumentation and Measurement*, 2012, Vol. 61, No. 5, pp. 1140–1148.

[15] M. Elad, *Sparse and Redundant Representations*, Springer Verlag, 2010.

[16] T. Hastie, R. Tibshirani, and M. Wainwright, *Statistical Learning with Sparsity: The Lasso and Generalizations*, Chapman & Hall/CRC, Boca Raton, Florida, 2015.

[17] E. T. Jaynes and G. L. Bretthorst, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.

[18] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[19] R. D. Luce and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.

[20] J. Ma and F.-X. Le Dimet, "Deblurring from highly incomplete measurements for remote sensing", *IEEE Transactions on Geosciences Remote Sensing*, 2009, Vol. 47, No. 3, pp. 792–802.

[21] B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman, San Francisco, California, 1983.

[22] L. McMackin, M. A. Herman, B. Chatterjee, and M. Weldon, "A high-resolution swir camera via compressed sensing", *Proceedings of SPIE*, 2012, Vol. 8353, No. 1, p. 8353-03.

[23] J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[24] B. Mittelstadt, C. Russell, and S. Wachter, "Explaining explanations in AI", *Proceedings of the 2019 ACM Fairness, Accounatbility, and Transparency Conference FAT'2019*, Atlanta, Georgia, January 29–31, 2019, pp. 279–288.

[25] J. Nash, "The bargaining problem", *Econometrica*, 1950, Vol. 18, No. 2, pp. 155–162.

[26] B. K. Natarajan, "Sparse approximate solutions to linear systems", *SIAM Journal on Computing*, 1995, Vol. 24, pp. 227–234.

[27] Hoang Phuong Nguyen, L. Bokati, and V. Kreinovich, "New (simplified) derivation of Nash's bargaining solution", *Journal of Advanced Computational Intelligence and Intelligent Informatics (JACIII)*, to appear.

[28] H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[29] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[30] C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Books, New York, 2016.

[31] R. Tibshirani, "Regression shrinkage and selection via the LASSO", *Journal of the Royal Statistical Society*, 1996, Vol. 58, No. 1, pp. 267–288.

[32] Y. Tsaig and D. Donoho, "Compressed sensing", *IEEE Transactions on Information Theory*, 2006, Vol. 52, No. 4, pp. 1289–1306.

[33] L. Xiao, J. Shao, L. Huang, and Z. Wei, "Compounded regularization and fast algorithm for compressive sensing deconvolution", *Proceedings of the 6th International Conference on Image Graphics*, 2011, pp. 616–621.

[34] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.