



SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS



UNIVERSITY OF MINNESOTA

Integrity Under Attack:

Challenges and responses in the math research community

Douglas N. Arnold

School of Mathematics, University of Minnesota
President, Society for Industrial and Applied Mathematics

2nd World Conference on Research Integrity

Singapore, 22 July 2010

A case of plagiarism

A SAMPLE APPROXIMATION APPROACH FOR OPTIMIZATION WITH PROBABILISTIC CONSTRAINTS*

JAMES LUEDTKE[†] AND SHABBIR AHMED[†]

Abstract. We study approximations of optimization problems with probabilistic constraints in which the original distribution of the underlying random vector is replaced with an empirical distribution obtained from a random sample. We show that such a sample approximation problem with a risk level larger than the required risk level will yield a lower bound to the true optimal value with probability approaching one exponentially fast. This leads to an a priori estimate of the sample size required to have high confidence that the sample approximation will yield a lower bound. We then provide conditions under which solving a sample approximation problem with a risk level smaller than the required risk level will yield feasible solutions to the original problem with high probability. Once again, we obtain a priori estimates on the sample size required to obtain high confidence that the sample approximation problem will yield a feasible solution to the original problem. Finally, we present numerical illustrations of how these results can be used to obtain feasible solutions and optimality bounds for optimization problems with probabilistic constraints.

Key words. probabilistic constraints, chance constraints, Monte Carlo, stochastic programming, large deviation

AMS subject classification. 90C15

DOI. 10.1137/070702928

FREE

[Sample Issue](#)

[Trial Access](#)

[Quick Search](#)

[Advanced Search](#)

Probabilistic Transportation Problem (PTP)

Sreenivas M.¹, Srinivas T.²

¹Alluri Institute of Management Sciences, Hunter Road, Hanamkonda-506001, Andhra Pradesh, India

²Department of Mathematics, Kakatiya University, Warangal-506009, Andhra Pradesh, India

Abstract

We study approximations of optimization problems with probabilistic constraints in which the original distribution of the underlying random vector is replaced with an empirical distribution obtained from a random sample. We show that such a sample approximation problem with risk level larger than the required risk level will yield a lower bound to the true optimal value with probability approaching one exponentially fast. This leads to an a priori estimate of the sample size required to have high confidence that the sample approximation will yield a lower bound. We then provide conditions under which solving a sample approximation problem with a risk level smaller than the required risk level will yield feasible solutions to the original problem with high probability. Once again, we obtain a priori estimates on the sample size required to obtain high confidence that the sample approximation problem will yield a feasible solution to the original problem. Finally, we present numerical illustrations of how these results can be used to obtain feasible solutions and optimality bounds for optimization problems with probabilistic constraints.

Start of the investigation

- 4/7/09: Author informs SIAM of duplicate abstract.
- 4/8/09: SIAM publisher writes letter of inquiry to publisher (RIP) and EIC next day.
- After 3 weeks w/o response writes again. Nothing from RIP, but EIC responds:

Start of the investigation

- 4/7/09: Author informs SIAM of duplicate abstract.
- 4/8/09: SIAM publisher writes letter of inquiry to publisher (RIP) and EIC next day.
- After 3 weeks w/o response writes again. Nothing from RIP, but EIC responds:

If you can contact RIP then you are doing better than any editor for any of their journals! There has been extensive correspondence among the editors and we are simply not sure that any of the journals exist. The entire publishing venture may have been a scam.

Start of the investigation

- 4/7/09: Author informs SIAM of duplicate abstract.
- 4/8/09: SIAM publisher writes letter of inquiry to publisher (RIP) and EIC next day.
- After 3 weeks w/o response writes again. Nothing from RIP, but EIC responds:

If you can contact RIP then you are doing better than any editor for any of their journals! There has been extensive correspondence among the editors and we are simply not sure that any of the journals exist. The entire publishing venture may have been a scam.

- Unable to raise a response from publisher or authors, SIAM undertakes a full investigation.

SIOPT preprint (posted to Optimization Online September 2007)

0.999 that the optimal solution is at most 0.6% less costly than this solution. Using $N = 5000$ (or greater), we obtain a feasible solution of the same cost, but a lower bound which states that with confidence at least 0.999 this feasible solution is optimal.

3.2 Probabilistic Transportation Problem

We next tested the sampling approach on a probabilistic version of the classical transportation problem, which we call the Probabilistic Transportation Problem (PTP). In this problem, we have a set of suppliers I and a set of customers D with $|D| = m$. The suppliers have limited capacity M_i for $i \in I$. There is a transportation cost c_{ij} for shipping a unit of product from supplier $i \in I$ to customer $j \in D$. The customer demands are random and are represented by a random vector \tilde{d} taking values in \mathbf{R}^m . We assume we

IJSS paper

numerical illustrations of how these results can be used to obtain feasible solutions and optimality bounds for optimization problems with probabilistic constraints.

1. Introduction

A set of suppliers I and a set of customers D with $|D| = m$. The suppliers have limited capacity M_i for $i \in I$. There is a transportation cost c_{ij} for shipping a unit of product from supplier $i \in I$ to customer $j \in D$. The customer demands are random and are represented by a random vector d taking values in \mathbf{R}^m . We assume we must choose the shipment quantities before the customer demands are known. We enforce the

Table 6: Lower bounds for low variance PTP sample problems with $\alpha = \epsilon = 0.05$.

N	LB with confidence at least:				Gap with confidence at least:			
	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
1000	1.9755	1.9757	1.9775	1.9782	1.55%	1.54%	1.45%	1.42%
3000	1.9879	1.9892	1.9892	1.9910	0.93%	0.87%	0.87%	0.78%
5000	1.9940	1.9943	1.9948	1.9951	0.63%	0.62%	0.59%	0.57%
7500	1.9954	1.9956	1.9959	1.9963	0.56%	0.55%	0.54%	0.52%
10000	1.9974	1.9977	1.9980	1.9981	0.46%	0.45%	0.43%	0.42%

Table 5 gives the characteristics of the solutions generated for the different values of α and N . We observe that as in the case of the PSC, the *average* cost of the feasible solutions obtained using $\alpha > 0$ is always less than the *minimum* cost of the feasible solutions obtained with $\alpha = 0$. However, for this instance, the minimum cost solution obtained using $\alpha = 0$ is not so significantly worse than the minimum cost solutions

Table 2 : Lower bounds for low variance PTP sample problems with $\alpha = \epsilon = 0.05$.

N	LB with confidence at least:				Gap with confidence at least:			
	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
1000	1.9755	1.9757	1.9775	1.9782	1.55%	1.54%	1.45%	1.42%
3000	1.9879	1.9892	1.9892	1.9910	0.93%	0.87%	0.87%	0.78%
5000	1.9940	1.9943	1.9948	1.9951	0.63%	0.62%	0.59%	0.57%
7500	1.9954	1.9956	1.9959	1.9963	0.56%	0.55%	0.54%	0.52%
10000	1.9974	1.9977	1.9980	1.9981	0.46%	0.45%	0.43%	0.42%

Table 1 gives the characteristics of the solutions generated for the different values of α and N . We observe that as in the case of the PSC, the average cost of the feasible solutions obtained using $\alpha > 0$ is always less than the minimum cost of the feasible solutions obtained with $\alpha = 0$. However, for this instance, the minimum cost solution obtained using $\alpha = 0$ is not so significantly worse than the minimum cost solutions since different values of α are giving between 0.46% and 0.56% more

which are not necessarily independent and identically distributed. For example, the use of variance reduction techniques such as Latin hypercube sampling or Quasi-Monte Carlo sampling may yield significantly faster convergence.

References

- [1] S. Ahmed and A. Shapiro. The sample average approximation method for stochastic programs with integer recourse. Preprint available at www.optimization-online.org, February 2002.
- [2] J. Atlason, M. A. Epelman, and S. G. Henderson. Call center staffing with simulation and cutting plane methods. *Annals of Operations Research*, 127:333–358, 2004.
- [3] J. Beasley. Or-library: Distributing test problems by electronic mail. *Journal of the Operational Research Society*, 41:1069–1072, 1990.
- [4] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23:760–805, 1998.

which are not necessarily independent and identically distributed. For example, the use of variance reduction techniques such as Latin hypercube sampling or Quasi-Monte Carlo sampling may yield significantly faster convergence.

References

- [1] Ahmed and Shapiro. The sample average approximation method for stochastic programs with integer recourse.
- [2] Atlason, Epelman, and Henderson. Call center staffing with simulation and cutting plane methods. *Annals of Operations Research*, 127:333–358, 2004.
- [3] Beasley. Or-library: Distributing test problems by electronic mail. *Journal of the Operational Research Society*, 41:1069–1072, 1990.
- [4] Ben-Tal and Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23:760–805, 1998.

The investigation proceeds

- We discovered about half a dozen cases of plagiarism by these same two authors, some known, some new. We informed the journals, resulting in several retractions and one cancelled publication.
- On 8/25/09 we wrote to the administration of each author's institution, and then heard from both authors.

The investigation proceeds

- We discovered about half a dozen cases of plagiarism by these same two authors, some known, some new. We informed the journals, resulting in several retractions and one cancelled publication.
- On 8/25/09 we wrote to the administration of each author's institution, and then heard from both authors.

I corresponding author, M. Sreenivas, take this humble opportunity to request you that the article titled Probabilistic Transportation Problem (PTP), appeared in Volume 3, Issue 1 of International Journal of Statistics and Systems, I ashamed for the situation and I have been stopped such type of submissions to the Journals. The co-author, which was mentioned in the article, is no way related to this matter, as it is not in the notice of him. It was happened without consulting the person. He is completely unaware of the article.

[M. Sreenivas]

The investigation proceeds

- We discovered about half a dozen cases of plagiarism by these same two authors, some known, some new. We informed the journals, resulting in several retractions and one cancelled publication.
- On 8/25/09 we wrote to the administration of each author's institution, and then heard from both authors.

I corresponding author, M. Sreenivas, take this humble opportunity to request you that the article titled Probabilistic Transportation Problem (PTP), appeared in Volume 3, Issue 1 of International Journal of Statistics and Systems, I ashamed for the situation and I have been stopped such type of submissions to the Journals. The co-author, which was mentioned in the article, is no way related to this matter, as it is not in the notice of him. It was happened without consulting the person. He is completely unaware of the article.

[M. Sreenivas]

Recently. . . I came to know that some bad things happened in 2007 without my knowledge and my name was included in the papers of M.Sreenivas, a student of O.R. and working in a private college. . . he corresponded every thing by including my name but without my notice. . . I have worked in Algebra but not in O.R. and I had no idea about O.R.

[T. Srinivas]



Highlights : Department of Mathematics

[Dept.Profile](#)[Faculty](#)[Syllabus](#)

Student profile according to programmes of study, gender, region etc

The ratio of Male/Female students who are joining.

10. Transportation Management & Services by T. Srinivas and M. Srinivas, [Kakatiya Business Review \(KBR\)](#), Vol. 3, No.1, September 2006, PP: 121-128, Warangal.
11. Effectiveness of Distribution Network by T. Srinivas and M. Srinivas, [International Journal of Information Systems and Supply Chain Management \(IJISSCM\)](#), Vol.1, No.1, January-March 2008, PP: 80-86, USA.
12. A 7 Step Approach for Transportation Optimization by T. Srinivas and M. Srinivas, [Pratibimba, Vol. 8, Issue: 1](#), Jan-June 2008, PP: 67-69, Bhubaneshwar.
13. Transportation Applications of Neural Networks by T. Srinivas and M. Srinivas, [International Journal of Computing and Applications \(IJCA\) Serials Publications](#), Vol.2, No.2, Dec 2007, PP: 157-162, New Delhi.
14. Modeling the Transportation Problem As a Flow Problem by T. Srinivas and M. Srinivas, Proceedings of the first International Conference on Emerging Technologies & Applications in Engineering, Technology & Sciences (ICETAETS – 2008), Vol. III, PP: 2616-2619, Gujarat.
15. The Role of Transportation in Logistics Chain by T. Srinivas and M. Srinivas, [Indian Journal of Mathematics and Mathematical Sciences \(IJMMS\) Serials Publications](#), Vol.4, No.2, June 2008, PP: 75-82, (in press), New Delhi.
16. . T.Srinivas and M. Srinivas : Transportation: More for less criteria, International Journal of Statistics and Management Svsstem, [Serials Publications](#), Vol. 3, No. 1–2, Jan, 2008. (in press), New Delhi.

Final report on the investigation

- 10/21/09: After more than 6 months of investigation and more than 300 emails, I wrote up a report (6 pages, 16 attachments) and we posted it on SIAM's web.
- I informed the authors, their institutions, and the entire editorial board of IJSS.
- A large fraction of the editorial board resigned, but their names remain posted.
- 12/17/09: The publisher responded to our initial inquiry, 8 months later.

Final report on the investigation

- 10/21/09: After more than 6 months of investigation and more than 300 emails, I wrote up a report (6 pages, 16 attachments) and we posted it on SIAM's web.
- I informed the authors, their institutions, and the entire editorial board of IJSS.
- A large fraction of the editorial board resigned, but their names remain posted.
- 12/17/09: The publisher responded to our initial inquiry, 8 months later.

Dear Dr. David K Marshall,^{←SIAM Publisher}

We have removed that article from our website and inform authors.

With regards,

Viveka Nand, Publication Department

Final report on the investigation

- 10/21/09: After more than 6 months of investigation and more than 300 emails, I wrote up a report (6 pages, 16 attachments) and we posted it on SIAM's web.
- I informed the authors, their institutions, and the entire editorial board of IJSS.
- A large fraction of the editorial board resigned, but their names remain posted.
- 12/17/09: The publisher responded to our initial inquiry, 8 months later.

Dear Dr. David K Marshall,
We have removed that article from our website and inform authors.
With regards,
Viveka Nand, Publication Department

← SIAM Publisher

- IJSS increased its frequency from 2 to 4 issues a year.

- WorldCat lists 116 libraries which subscribe to IJSS. Subscriptions are \$380 for print and online, \$360 online only.
- RIP publishes 131 math journals in math, science, and engineering.
- IJSS has mandatory page charges of \$20/page.
- RIP is far from alone. Unethical, predatory publishers are becoming very common.
- Even this super-clear case was very time-consuming to investigate, and could have easily gone unnoticed. More subtle ethical breaches are much more difficult.

The Impact Factor in Applied Mathematics

Journal Impact Factor: “Mathematics, Applied”

Journal	2009	2008	2007	2006	2005	2004
CPAM	3 (2.7)	2 (3.7)	2 (2.7)	4	7	5
IJNSNS	1 (5.3)	1 (8.9)	1 (5.1)	1	2	7
SIREV	2 (3.4)	3 (2.8)	3 (2.5)	3	1	1

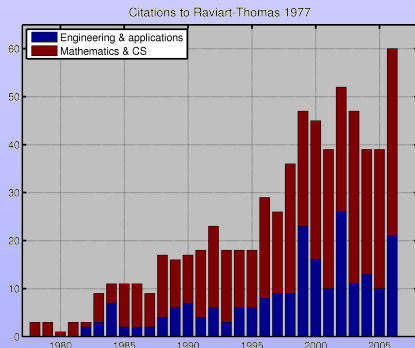
- Communications on Pure and Applied Mathematics
- International Journal on Nonlinear Sciences and Numerical Simulation
- SIAM Review

Journal	articles in 2006–7	2008 cites to these	impact factor
CPAM	93	343	3.7
IJNSNS	163	1453	8.9
SIREV	46	129	2.8

Obvious flaws in the IF as a **journal quality** metric

- IF captures only a tiny bit of the citation information of a journal.
- Journal quality is not captured even by complete citation information.
- The two-year window is *much* too small for mathematics.
- IF reflects the citation culture of the discipline and subdiscipline.
- Not all citing journals are in the JCR database.
- There are errors in the JCR data.

Etc., etc.



Impact factor manipulation

There is mounting concern that attempts to manipulate impact factors are harming scientific research.

– Wall Street Journal, June 5, 2006

The EIC of *J. Gerontology A* authored and published a review article every January focusing on the preceding two years. E.g., in January 2004, his article *The Top 10 Hot Topics in Aging* had 277 references, 195 of which were to *J. Gerontology A* in 2002 or 2003.

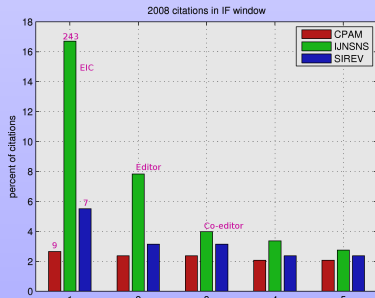
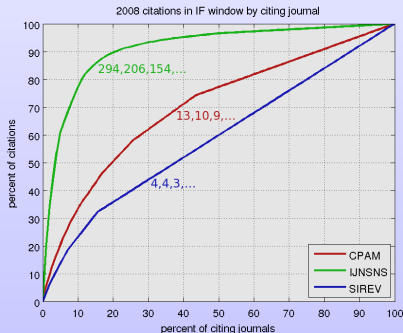
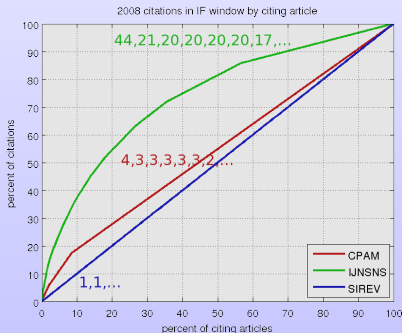
In order to preserve a high quality of our journals, the Editorial Board has established the following rules:

a) is advisable for each accepted paper to contain citations to articles published during 2006-2008 in our journals. . .

– *Instructions to Authors*, Balkan Journal of Geometry and Its Applications

The journal contributing the most citations to the IJNSNS 2008 IF was a conference proceedings, organized and edited by the IJNSNS EIC. It had 366 citations to IJNSNS (294 in the IF window) and 353 citations of the EIC himself.

Some citation distribution analysis



2008 IF with different windows

Journal	window 2006-7	window 2004-5	window 2002-3
CPAM	3.7	4.6	2.9
IJNSNS	8.9	3.8	0.4
SIREV	2.8	4.5	20.8

Metric-based ranking of individuals



TRACKING TRENDS & PERFORMANCE IN BASIC RESEARCH

Interviews Analyses Data & Rankings

2008 : April 2008 - Author Commentaries : Professor JH He

AUTHOR COMMENTARIES - 2008

April 2008



Professor JH He
A Featured Scientist from **Essential Science IndicatorsSM**

According to a recent analysis of **Essential Science Indicators** from **Thomson Scientific**, Professor Ji-Huan He has been named a **Rising Star** in the field of Computer Science. His citation record in this field includes 21 papers cited a total of 306 times between January 1, 1997 and December 31, 2007. He also has 25 papers cited a total of 881 times in Engineering, and 16 papers cited a total of 87 times in Materials Science. His citation record in the Web of Science[®] includes 137 papers cited a total of 3,193 times to date.

Thomson–Reuters named IJNSNS EIC:

- “Rising Star” in Computer Science in 2008
- “New Hot Paper” in Physics in 2008
- “Hot Paper” in Mathematics in 2008
- “Fast Breaking Paper” in Engineering in 2008
- “Hottest Research of 2007–08” (1 of 13 in all fields of science)
- “Hottest Research of 2009” (1 of 12)

Food for thought

- There can be little doubt that the IF is highly flawed as an indicator of journal, paper, or author quality.
- But the urge to have a simple metric is irresistible to many.

We must stop the avalanche of low-quality research. . . Several fixes come to mind: . . . Make more use of citation and journal “impact factors,” from Thomson ISI. . . add [journal Impact factors] to a researcher’s publication record.

— Commentary in The Chronicle of Higher Education, June 13, 2010

Food for thought

- There can be little doubt that the IF is highly flawed as an indicator of journal, paper, or author quality.
- But the urge to have a simple metric is irresistible to many.

We must stop the avalanche of low-quality research. . . Several fixes come to mind: . . . Make more use of citation and journal “impact factors,” from Thomson ISI. . . add [journal Impact factors] to a researcher’s publication record.

— Commentary in The Chronicle of Higher Education, June 13, 2010

- I have shown gross anomalies, but even these were extremely time-consuming to discover and document, and required expertise.
- How many journals have manipulated themselves up the IF list, trodding over journals run with more integrity?
- The problem cannot be solved by adjusting the formulas, using eigenvalues, statistics, . . .

Food for thought

- There can be little doubt that the IF is highly flawed as an indicator of journal, paper, or author quality.
- But the urge to have a simple metric is irresistible to many.

We must stop the avalanche of low-quality research. . . Several fixes come to mind: . . . Make more use of citation and journal “impact factors,” from Thomson ISI. . . add [journal Impact factors] to a researcher’s publication record.

— Commentary in The Chronicle of Higher Education, June 13, 2010

- I have shown gross anomalies, but even these were extremely time-consuming to discover and document, and required expertise.
- How many journals have manipulated themselves up the IF list, trodding over journals run with more integrity?
- The problem cannot be solved by adjusting the formulas, using eigenvalues, statistics, . . .

Arnold’s Theorem: Any “objective” indicator based largely on recent citation data, is easily gamed.

Research metrics

If you count tree rings to estimate the age, the trees don't start producing extra rings. Counting papers or citations is different.



REFERENCES

- [1] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus applications. *Acta Numer.*, 15:1–155, 2006.
- [2] I. Babuska. The finite element method with Lagrangian multipliers. *Nucl. Energy, Part C: J. Plasma Physics*, 30:15–32, 1982.
- [3] D. Boffi, F. Brezzi, and M. Fortin. Finite elements for the Stokes problem: *Compatibility Conditions, and Applications*, volume 1939 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2008. Lectures given at the C.I.M.E. Summer School, July 1, 2006, Edited by D. Boffi and L. Gastaldi.
- [4] D. Boffi, F. Brezzi, and L. Gastaldi. On the problem of spurious eigenvalues for linear elliptic problems in mixed form. *Math. Comp.*, 69(229):121–140, 1997.
- [5] D. Boffi, R. G. Duran, and L. Gastaldi. A remark on spurious eigenvalues for the Stokes problem. *Math. Comp.*, 72(243):107–114, 1999.
- [6] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*. Springer, New York, third edition, 2008.
- [7] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems for elliptic problems with general Lagrangian multipliers. *Rev. Française Automat. Informat. Recherche Opérationnelle*, 12(2):129–151, 1974.
- [8] F. Brezzi, J. Douglas, Jr., and L. D. Marini. Two families of mixed elliptic problems. *Numer. Math.*, 47(2):217–235, 1985.
- [9] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Computational Mathematics. Springer-Verlag, New York, 1991.
- [10] D. Chapelle and K.-J. Bathe. The inf-sup test. *Comput. & Structures*, 17(1):37–45, 1983.

The discipline which throws the most light on research metrics is not statistics, but game theory.

A proposal

A proposal to the International Mathematical Union and the International Council on Industrial and Applied Mathematics:

IMU and ICIAM should convene a large, broad committee of mathematical experts of the highest level and integrity to produce a coarse rating of mathematics journals.

- This will provide an easily applied measure of journal quality.
- Useful for librarians, authors, editors, publishers, evaluators, . . .
- It will *not* measure author quality or paper quality!
- Safeguards against conflict-of-interest.

Other math community actions

- *Citation Statistics*, a report from IMU, ICIAM, and IMS, was published in 2008.
- My 2010 editorial *Integrity Under Attack* has been translated into 4 other languages and widely circulated.
- A new IMU document on *Best Practices for Mathematics Journals* is nearly finalized.
- A panel discussion on research metrics will be held at the International Congress of Mathematicians in India in August.