

Math Problems Collection

2020年12月5日 11:42

This document collects some interesting (stand-alone) math problems and their proofs. They are collected from literature study.

Version: Dec 5, 2020

❖ Table of Contents

[\[Linear Programming\] Randomized Kaczmarz Algorithm](#)

[\[MATRIX\] Orthogonal Projection](#)

[\[Linear Algebra\] Vector Decomposition, Span and Null](#)

[\[STATS\] L1 and L2 regulation for feature selection](#)

❖ [Linear Programming] Randomized Kaczmarz Algorithm

- **Reference:**

T. Strohmer and R. Vershynin. "Randomized Kaczmarz Algorithm with Exponential Convergence". Journal of Fourier Analysis and Applications 15.2 (2009).

<https://people.eecs.berkeley.edu/~brecht/cs294docs/week1/09.Strohmer.pdf>

- Classical Kaczmarz's method runs through a cyclic iteration

- Random version:

Algorithm 1 (Random Kaczmarz algorithm). *Let $Ax = b$ be a linear system of equations as in (1) and let x_0 be arbitrary initial approximation to the solution of (1). For $k = 0, 1, \dots$ compute*

$$\blacksquare \quad x_{k+1} = x_k + \frac{b_{r(i)} - \langle a_{r(i)}, x_k \rangle}{\|a_{r(i)}\|_2^2} a_{r(i)}, \quad (4)$$

where $r(i)$ is chosen from the set $\{1, 2, \dots, m\}$ at random, with probability proportional to $\|a_{r(i)}\|_2^2$.

- Proof is through comparing the distance to optimal solution across iterations. Using orthogonal projection, distance reduction is bounded by an upper limit, which is related to matrix conditional number.

❖ [MATRIX] Orthogonal Projection

- Please see <https://www.math.utk.edu/~mengesha/teaching/Math2025/Lecture8.pdf>

- <https://textbooks.math.gatech.edu/ila/projections.html>

❖ [Linear Algebra] Vector Decomposition, Span and Null

- A given vector can be decomposed into a sum of two vectors, one is in span space and the other in null space
- From {T. Strohmer and R. Vershynin. "Randomized Kaczmarz Algorithm with Exponential Convergence"}
direction, so we shall decompose $\mathbb{R}^n = \text{span}(x_0) \oplus (x_0)^\perp$, writing every vector $x \in \mathbb{R}^n$ as
 - $x = x' \cdot x_0 + x''$, where $x' \in \mathbb{R}$, $x'' \in (x_0)^\perp$.
- Other links include:
 - <https://math.stackexchange.com/questions/1767601/why-any-vector-can-be-wriiten-as-the-sum-of-two-components-in-the-row-space-and>
 - <https://www.math.utah.edu/~zwick/Classes/Fall2012/2270/Lectures/Lecture21.pdf>
 - An introduction to vector space:
<https://www2.stat.duke.edu/courses/Spring06/sta244/Handouts/vectorspaces.pdf>

❖ [STATS] L1 and L2 regulation for feature selection

- L1 regulation is usually used for feature selection
 - <https://towardsdatascience.com/ridge-and-lasso-regression-a-complete-guide-with-python-scikit-learn-e20e34bcbf0b>
 - <http://ethen8181.github.io/machine-learning/regularization/regularization.html>