# Blendenpik: Randomized Least Squares

Project - Computational Linear Algebra, EPFL, Spring 22

Moritz Waldleben

May 2022

## 1 Introduction

Blendenpik algorithm [1] is a randomized algorithm to solve a linear least squares problem. The iterative algorithm used for solving this minimization problem is called LSQR. Blendenpik constructs a random projection preconditioner for the LSQR algorithm which convergences in fiewer steps than an unpreconditioned LSQR.

The Blendenpik algorithm can beat the standard LAPACK implementation for very large matrices. In this report we will discuss the main concepts. In this report we will investigate the method proposed in the paper.

Instead of the LSQR algorithm we will use the minimum residual method (MINRES) with preconditioning. MINRES will be applied to the normal equation of the least squares problem. A Matlab implementation is given and some illustrative results will be shown.

## 2 Motivation

Let us look at a large overdetermined system: $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, m >> n,$ and $\operatorname{rank}(A) = n$. The corresponding linear squares problem can be written as

$$\min_{x \in \mathbb{R}^n} \|A\mathbf{x} - \mathbf{b}\|_2$$

## 3 The uniform random sampling

To tackle uniform random sampling an important property for the matrix $A$ is coherence. Coherence basicly tells us how much does the solution of the system will depend one row. Formally coherence can be defined as

$$\mu(A) = \max_i \|U(i, :)\|_2^2,$$

where $\frac{n}{m} \leq \mu(A) \leq 1$.

From the above definition we can think of two extrem ewample cases:

1. Zero entry in $A$ will lead to a large row
2. A has only one nonzero entry in one of the colums

We will now prove bounds for the coherence number $\mu(A)$ for a general matrix $A \in \mathbb{R}m \times n$.

The following theorem gives us a probabilistic upper bound the condition number of $AR^{-1}$. This will later be usefull the ensure that the preconditioned iterative solver will converge in fiew iterations.

**Theorem 1** *blendenpik Let $A$ be an $m \times n$ full rank matrix, and let $\mathcal{S}$ be a random sampling operator that samples $r \geq n$ rows from $A$ uniformly. Let $\tau = C\sqrt{m\mu(A)\log(r)/r}$, where $C$ is some constant defined in the proof. Assume that $\delta^{-1}\tau < 1$. With probability of at least $1-\delta$, the sampled matrix $\mathcal{S}A$ is full rank, and if $\mathcal{S}A = QR$ is a reduced QR factorization of $\mathcal{S}A$, we have*

$$\kappa\left(AR^{-1}\right) \leq \frac{1+\delta^{-1}\tau}{1-\delta^{-1}\tau}$$

To prove the above we will state and use two theorems which correspond to theorem 1 in [2] and theorem 7 in [3].

**Theorem 2** *CUR Suppose that $l, m,$ and $n$ are positive integers such that $m \geq l \geq n$. Suppose further that $A$ is a full-rank $m \times n$ matrix, and that the SVD of $A$ is*

$$A_{m\times n} = U_{m\times n}\Sigma_{n\times n}V^*_{n\times n}.$$

*Suppose in addition that $T$ is an $l \times m$ matrix such that the $l \times n$ matrix $TU$ has full rank.*

*Then, there exist an $n \times n$ matrix $P$, and an $l \times n$ matrix $Q$ whose columns are orthonormal, such that*

$$T_{l\times m}A_{m\times n} = Q_{l\times n}P_{n\times n}.$$

*Furthermore, if $P$ is any $n \times n$ matrix, and $Q$ is any $l \times n$ matrix whose columns are orthonormal, such that $P$ and $Q$ satisfy Eq. 27, then the condition numbers of $AP^{-1}$ and $TU$ are equal.*

**Theorem 3** *Suppose $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times p}$, and $c \leq n$. Construct $C$ and $R$ with Algorithm 6, using the EXPECTED $(c)$ algorithm. If the sampling probabilities $\{p_i\}_{i=1}^n$ used by the algorithm are of the form (44) or (45), then*

$$\mathbf{E}\left[\|AB - CR\|_F\right] \leq \frac{1}{\sqrt{\beta c}}\|A\|_F\|B\|_F.$$

*If, in addition, $B = A^T$, then*

$$\mathbf{E}\left[\left\|AA^T - CC^T\right\|_2\right] \leq O(1)\sqrt{\frac{\log c}{\beta c}}\|A\|_F\|A\|_2.$$

Now we will prove theorem **??**.

# 4 The Blendenpik algorithm

As mentioned bevor the In this report we will primarly focus on the DCT transform. Additionally instead of the LSQR algorithm the the Minimal residual method (MINRES) will be used to solve the preconditioned least squares problem. To apply MINRES the normal equation of...

The grasp the idea idea of the algorithm a superficial pseudocode is given below.

# 5 Implementation

# 6 Numerical Experiments

# References


[1] Haim Avron and Petar Maymounkov. Blendenpik: Supercharging lapack's least-squares solver. *SIAM J. Scientific Computing*, 32:1217–1236, 01 2010.

[2] Petros Drineas, Michael Mahoney, and Senthilmurugan Muthukrishnan. Relative-error *cur* matrix decompositions. *SIAM Journal on Matrix Analysis and Applications*, 30:844–881, 05 2008.

[3] Vladimir Rokhlin and Mark Tygert. A fast randomized algorithm for overdetermined linear least-squares regression. *Proceedings of the National Academy of Sciences of the United States of America*, 105:13212–7, 09 2008.