Reducing Memory Access Latencies using Data Compression in Sparse, Iterative Linear Solvers

An All-College Thesis

College of Saint Benedict/Saint John's University

by Neil Lindquist

April 2018

Project Title: Reducing Memory Access Latencies using Data Compression in Sparse, Iterative Linear Solvers
Approved by:

Mike Heroux
Scientist in Residence

Robert Hesse
Associate Professor of Math

Jeremy Iverson
Assistant Professor of Computer Science

Bret Benesh
Chair, Department of Mathematics

Imad Rahal
Chair, Department of Computer Science

Director, All College Thesis Program

Abstract

Contents

1 Introduction

2 Background

2.1 Conjugate Gradient

Conjugate Gradient is the iterative solver used by HPCG [1]. Symmetric, positive definite matrices will guarantee the converge of Conjugate Gradient to the correct solution within n iterations when using exact algebra [2]. As an iterative method, Conjugate Gradient can provide a solution, $\vec{x}$, where $\left\|\mathbf{A}\vec{x} - \vec{b}\right\|$ is within some tolerance, after significantly fewer than n iterations, allowing it to find solutions to problems where even n iterations is infeasible [3].

To understand the Conjugate Gradient, first consider the quadratic form of $\mathbf{A}\vec{x} = \vec{b}$. The quadratic form is a function $f : \mathbb{R}^n \to \mathbb{R}$ where

$$f(\vec{x}) = \frac{1}{2}\vec{x} \cdot (\mathbf{A}\vec{x}) - \vec{b} \cdot \vec{x} \tag{1}$$

for some $c \in \mathbb{R}$. Note that

$$\begin{aligned}
\nabla f\left(\vec{x}\right) &= \nabla \left(\frac{1}{2}\vec{x} \cdot (\mathbf{A}\vec{x}) - \vec{b} \cdot \vec{x}\right) \\
&= \frac{1}{2}\nabla\left(\vec{x} \cdot (\mathbf{A}\vec{x})\right) - \nabla\left(\vec{b} \cdot \vec{x}\right) \\
&= \frac{1}{2}\left(\mathbf{A}\vec{x} + \mathbf{A}^T\vec{x}\right) - \vec{b}
\end{aligned}$$

Then, when $\mathbf{A}$ is symmetric,

$$\frac{\partial f(\vec{x})}{\partial \vec{x}} = \mathbf{A}\vec{x} - \vec{b}$$

So, the solution to $\mathbf{A}\vec{x} = \vec{b}$ is the sole critical point of f. Since $\mathbf{A}$ is the Hessian matrix of f at the point, if $\mathbf{A}$ is positive definite, then that critical point is a minimum. Thus, if $\mathbf{A}$ is a symmetric, positive definite matrix, then the minimum of f is the solution to $\mathbf{A}\vec{x} = \vec{b}$ [3].

The Method of Steepest Decent is similar to Conjugate Gradient, but simpler to initially understand. The method takes an initial $\vec{x}_0$ and repeatedly computes improved points, $\vec{x}_1, \vec{x}_2, \ldots$, until reaching a point close enough to the minimum of Equation 1. Because the gradient at a point is the direction of maximal increase, $\vec{x}_{i+1} = \vec{x}_i + \alpha\vec{r}_i$ for some $\alpha > 0$ and where $\vec{r}_i = -\nabla f\left(\vec{x}_i\right) = \vec{b} - \mathbf{A}\vec{x}_i$ is the residual of $\vec{x}_i$. The value for α that minimizes $f\left(\vec{x}_{i+1}\right)$ is when

$$\begin{aligned}
0 &= \frac{\partial f\left(\vec{x}_{i+1}\right)}{\partial \alpha} \\
&= \frac{\partial f\left(\vec{x}_i + \alpha \vec{r}_i\right)}{\partial \alpha} \\
&= \nabla f\left(\vec{x}_i + \alpha \vec{r}_i\right) \cdot \vec{r}_i \\
&= \left(\mathbf{A}\left(\vec{x}_i + \alpha \vec{r}_i\right) - \vec{b}\right) \cdot \vec{r}_i \\
&= \left(\mathbf{A}\vec{x}_i - \vec{b}\right) \cdot \vec{r}_i + \alpha \mathbf{A}\vec{r}_i \cdot \vec{r}_i \\
-\alpha \mathbf{A}\vec{r}_i \cdot \vec{r}_i &= \left(\mathbf{A}\vec{x}_i - \vec{b}\right) \cdot \vec{r}_i \\
\alpha &= \frac{\vec{r}_i \cdot \vec{r}_i}{\mathbf{A}\vec{r}_i \cdot \vec{r}_i}.
\end{aligned}$$

The resulting steps for the Method of Steepest Decent are

$$\begin{aligned}
\vec{r}_i &= \vec{b} - \mathbf{A}\vec{x}_i \\
\alpha &= \frac{\vec{r}_i \cdot \vec{r}_i}{\mathbf{A}\vec{r}_i \cdot \vec{r}_i} \\
\vec{x}_{i+1} &= \vec{x}_i + \alpha \vec{r}_i
\end{aligned}$$

until $\|\vec{r}_i\|$ is less than some tolerance [3].

Example 1.

3 Test Results

4 Conclusions and Future Work

5 References

[1] Jack Dongarra, Michael Heroux, and Piotr Luszczek. Hpcg benchmark: a new metric for ranking high performance computing systems. Technical Report UT-EECS-15-736, Electrical Engineering and Computer Sciente Department, Knoxville, Tennessee, November 2015.

[2] Y. Saad. *Iterative Methods for Sparse Linear Systems.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2nd edition, 2003.

[3] Jonathan R Shewchuk. An introduction to the conjugate gradient method without the agonizing pain. Technical report, Carnegie Mellon University, Pittsburgh, PA, USA, 1994.