Combinatorial Optimization of Matrix-Vector Multiplication

Michael Wolf, University of Illinois

Abstract: Work by Kirby, et al. (2006) showed that combinatorial optimization of matrix/vector multiplication could lead to faster evaluation of finite element stiffness matrices. Using relationships between rows, an efficient set of operations can be generated to perform matrix-vector multiplication. My improved graph model of this problem solves this combinatorial optimization problem optimally for binary row relationships. I extend the representation by using hypergraphs to model more complicated row relationships, expressing an n-row relationship with an n-vertex hyperedge. My initial greedy algorithm for this hypergraph model has yielded significantly better results than the graph model for many matrices.

Motivation

- Reducing redundant operations in building stiffness matrices
  - Generate code to optimize construction of local stiffness matrices
  - Many local stiffness matrices built
  - Reuse optimized code when problem is rerun
- Finite element "Compilers"
  - FIAT (automates generation of FEs)
  - FFC (variational forms -> code for evaluation)
- Following on work by Kirby, et al., Texas Tech, University of Chicago
  - Optimization of FFC generated code
  - FErari

Matrix-Vector Multiplication

Kirby, et al. showed that each element of FE local stiffness matrix can be calculated by the following Frobenius product:

$$S^e_{i,j} = K_{i,j} : G^e = \sum_m^2 \sum_n^2 G^e_{m,n} K_{i,j,m,n}$$

Flattening the stiffness matrix into a vector, we get a matrix-vector multiplication operation (y=Ax) with the following inner products for each row:

$$y_{ni+j} = S^e_{i,j} = A_{(ni+j,*)}x$$

For the 2D Laplace equation, for example, we get the inner products of the following vectors:

$$A^T_{(ni+j,*)} = \begin{bmatrix} \left(\frac{\partial \phi_i}{\partial r}, \frac{\partial \phi_j}{\partial r}\right)_{\hat{e}} \\ \left(\frac{\partial \phi_i}{\partial r}, \frac{\partial \phi_j}{\partial s}\right)_{\hat{e}} \\ \left(\frac{\partial \phi_i}{\partial s}, \frac{\partial \phi_j}{\partial r}\right)_{\hat{e}} \\ \left(\frac{\partial \phi_i}{\partial s}, \frac{\partial \phi_j}{\partial s}\right)_{\hat{e}} \end{bmatrix} \qquad x = \det(J) \begin{bmatrix} \frac{\partial r}{\partial x}\frac{\partial r}{\partial x} + \frac{\partial r}{\partial y}\frac{\partial r}{\partial y} \\ \frac{\partial r}{\partial x}\frac{\partial s}{\partial x} + \frac{\partial r}{\partial y}\frac{\partial s}{\partial y} \\ \frac{\partial s}{\partial x}\frac{\partial r}{\partial x} + \frac{\partial s}{\partial y}\frac{\partial r}{\partial y} \\ \frac{\partial s}{\partial x}\frac{\partial s}{\partial x} + \frac{\partial s}{\partial y}\frac{\partial s}{\partial y} \end{bmatrix}$$

Element independent

Element dependent

Optimization Problem

Objective: To minimize the number of multiply add pairs (MAPs) in matrix/vector multiplication

e.g. $r_2 = \alpha r_1 \Rightarrow y_2 = \alpha y_1$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} r_1^T \\ \alpha r_1^T \\ \vdots \\ r_m^T \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} r_1^T x \\ \alpha y_1 \\ \vdots \\ r_m^T x \end{bmatrix}$$

Possible Optimizations

Number of Nonzeros

$y_3 = 0$ 0 MAPs

$y_1 = 2x_1$ 1 MAP

$y_2 = 2x_1 + 2x_2$ 2 MAPs

$y_4 = 2x_1 + 2x_2 + 2x_3$ 3 MAPs

Scalar Multiples

$y_4 = y_2$ 0 MAPs

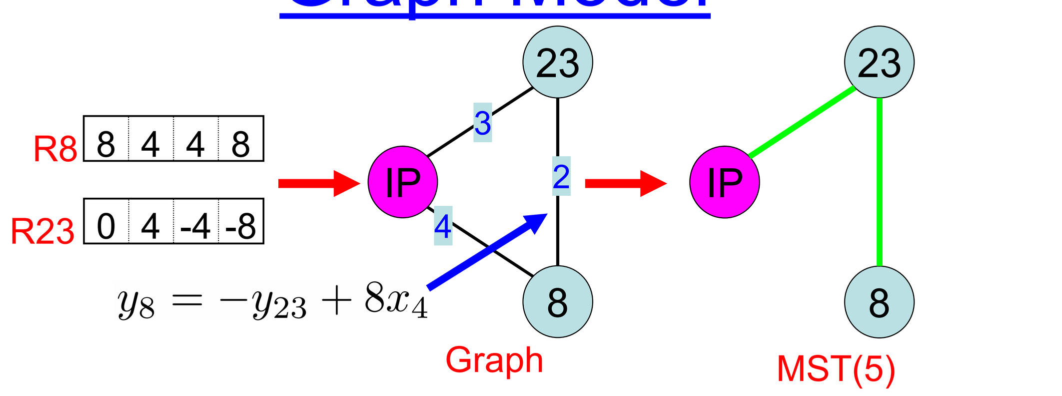
$y_2 = 4y_1$ 1 MAP

Partial Scalar Multiple $y_2 = 4y_4 + 4x_4$ 2 MAPs

Linear Combinations

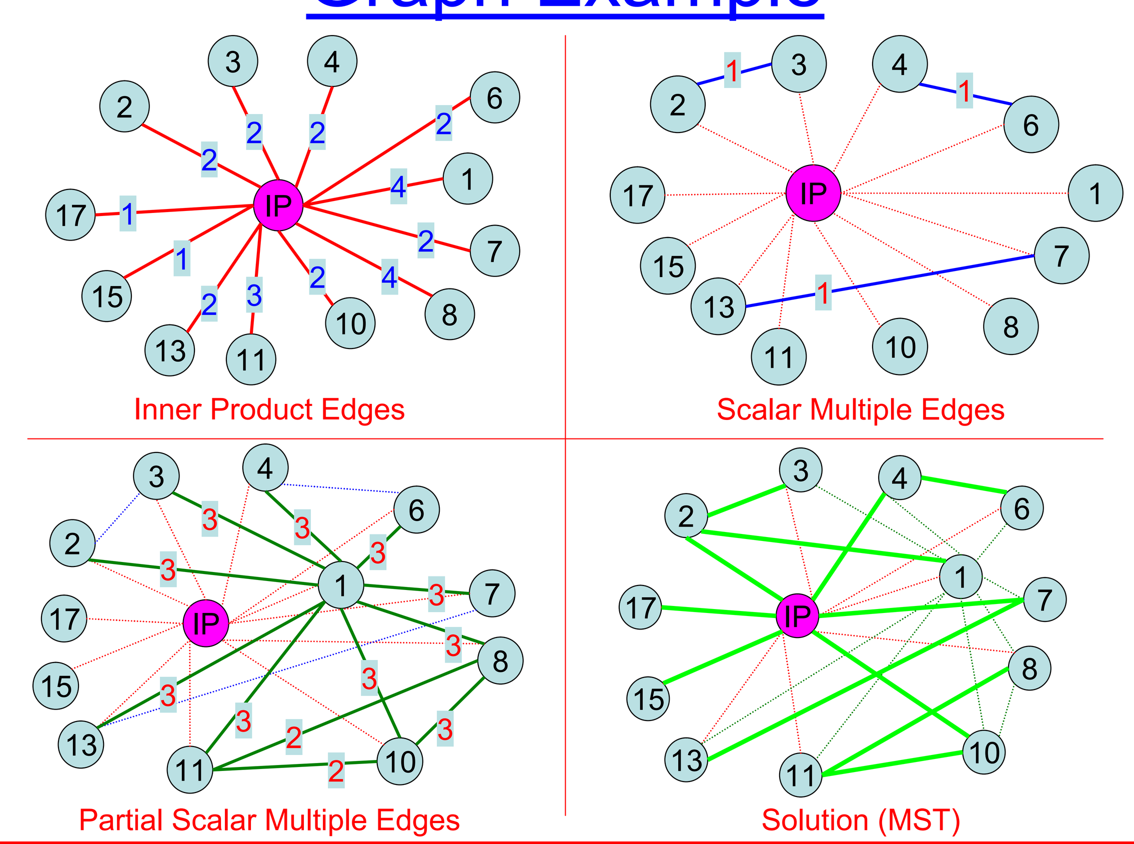
$y_2 = 4y_3 + 4y_4$ 2 MAPs

Graph Model



- Relationships between rows represented by edges between corresponding vertices
- Additional edges connect each vertex to IP vertex, representing inner product
  - FErari does not include this relationship in graph model
- Edge weights are MAP cost for particular operation
- Edges in minimum spanning tree (MST) represent operations to optimally compute (with these operations) matrix-vector multiplication
- In my work, I used Prim's algorithm to calculate MST

Graph Example



Graph Model Results

2D Laplace FErari Matrices, Matvec Costs

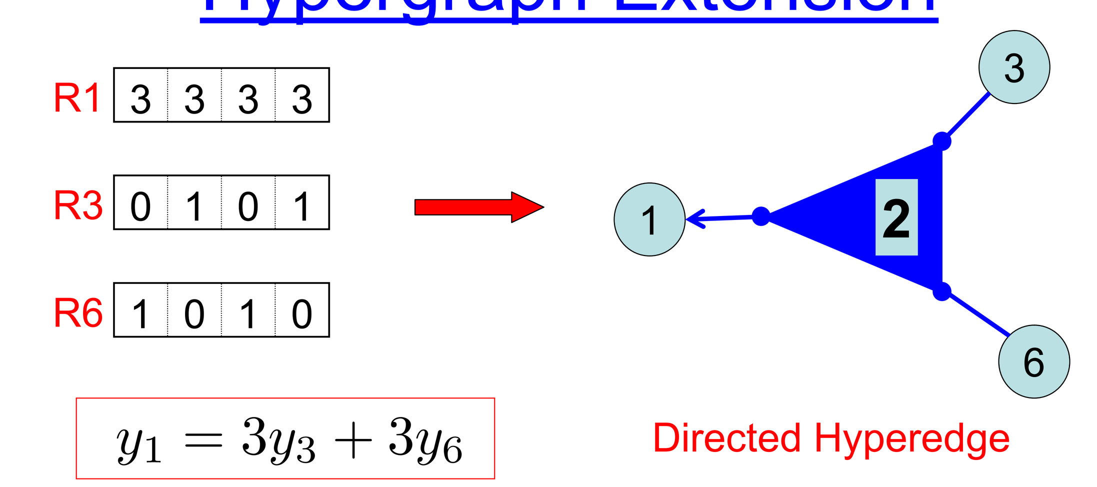
Order	Unoptimized MAPs	Nonzero MAPs	FErari MAPs	My MAPs
1	18	10	7	7
2	63	34	15	14*
3	165	108	45	43
4	360	292	176	152
5	693	589	443	366
6	1218	1070	867	678

3D Laplace FErari Matrices, Matvec Costs

Order	Unoptimized MAPs	Nonzero MAPs	FErari MAPs	My MAPs
1	60	21	-	17
2	330	177	94	79
3	1260	789	362	342
4	3780	2586	1118	1049
5	9576	7125	-	3567

*optimal

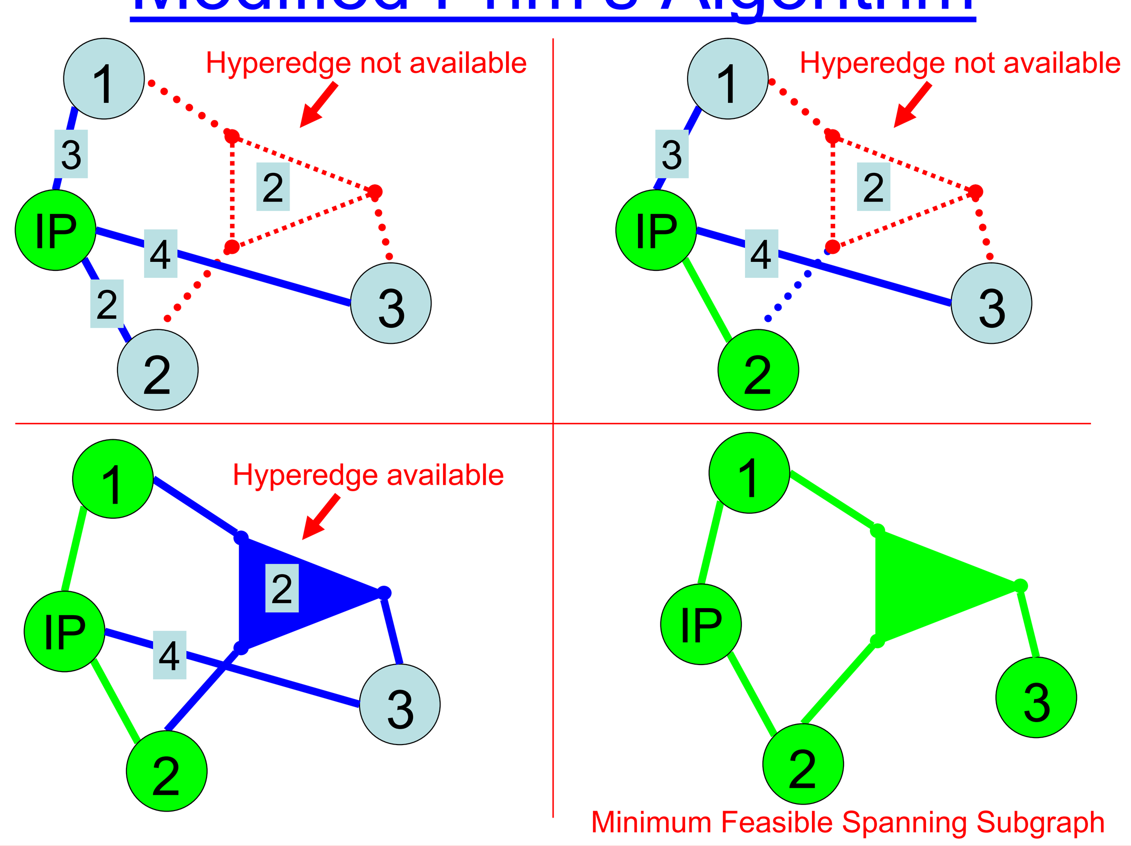
Hypergraph Extension



First Pass: Greedy Algorithm

- Modified Prim's algorithm to include hyperedges
- Polynomial time algorithm
- Solution no longer a tree
- No guarantee of optimum solution for hypergraph

Modified Prim's Algorithm



Hypergraph Model Results

2D Laplace FErari Matrices, Matvec Costs

Order	Unoptimized MAPs	Nonzero MAPs	Graph MAPs	HGraph MAPs
1	18	10	7	6*
2	63	34	14*	14*
3	165	108	43	43
4	360	292	152	150
5	693	589	366	363
6	1218	1070	678	678

3D Laplace FErari Matrices,Matvec Costs

Order	Unoptimized MAPs	Nonzero MAPs	FErari Geom. MAPs	My Graph MAPs	My HGraph MAPs
1	60	21	–	17	14*
2	330	177	105	79	68
3	1260	789	327	342	296
4	3780	2586	1045	1049	833
5	9576	7125	–	3567	3250

*optimal

Matrix-Vector Multiplication Timings

2D Laplace FErari Matrices

Order	Original Time (s)	Optimized Time (s)
1	2.16×10^{-3}	1.80×10^{-4}
2	7.74×10^{-3}	6.95×10^{-4}
3	1.94×10^{-2}	1.87×10^{-3}
4	4.23×10^{-2}	5.20×10^{-3}
5	8.09×10^{-2}	1.02×10^{-2}
6	1.53×10^{-1}	1.85×10^{-2}

3D Laplace FErari Matrices

Order	Original Time (s)	Optimized Time (s)
1	7.09×10^{-3}	3.65×10^{-4}
2	3.70×10^{-2}	1.68×10^{-3}
3	1.72×10^{-1}	9.62×10^{-3}
4	5.96×10^{-1}	3.15×10^{-2}

- 10,000 matrix vector multiplications per timing

Future Work

- Larger size hyperedges
  - Have implemented 3 and 4 vertex hyperedges
  - Higher degree perhaps useful for 3D FE problems
- More nearly optimal and optimal solution methods

Acknowledgements

- Robert C. Kirby – assistance with FErari and discussions about the problem in general
- Erik G. Boman – collaboration over the summer
- Michael T. Heath – my thesis advisor
- DOE CSGF fellowship support, DE-FG02-97ER25308

References

Kirby, R.C. "Combinatorial Structure in Evaluating Finite Element Operators," CSC'07, Monday 10:45 AM.

Kirby, R.C. and L. R. Scott. "Geometric Optimization of the Evaluation of Finite Element Matrices. To appear in SIAM Journal on Scientific Computing.

Kirby, R. C. (2006). FErari (version 0.1.0), University of Chicago.

Kirby, R. C., et al. (2006). "Topological optimization of the evaluation of finite element matrices." SIAM Journal on Scientific Computing 28: 224-240.

Kirby, R. C., et al. (2005). "Optimizing the Evaluation of Finite Element Matrices." SIAM Journal on Scientific Computing 27: 741-758.