A Multiscale Technique for Decomposing Sparse Matrices

Lek-Heng Lim

(Joint work with Gunnar Carlsson and Vin de Silva)

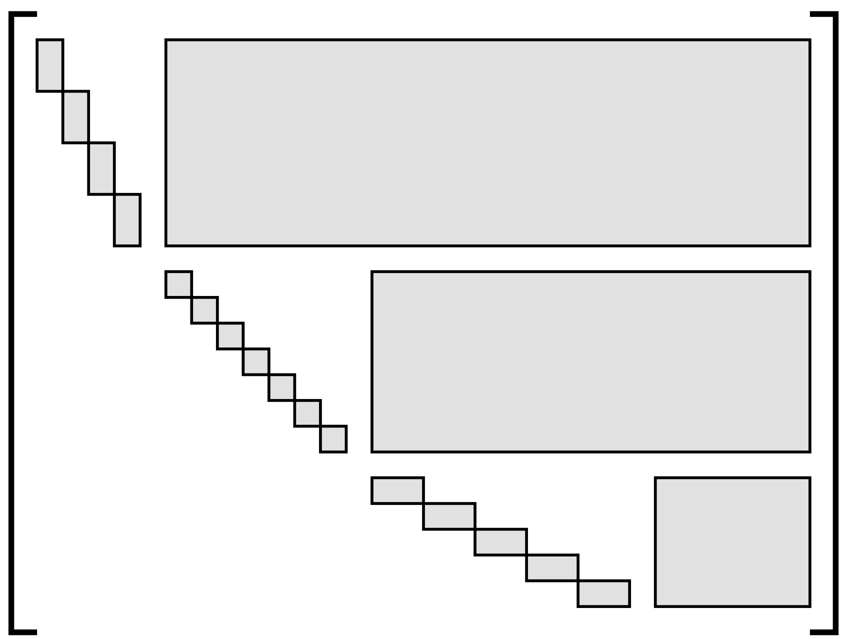
Program in Scientific Computing/Computational Mathematics and the Department of Mathematics, Stanford University

Thanks: NSF Grant DMS 01-01364, SIAM Student Travel Fund, Gene Golub

Basic Idea I: Geometric Sparsity

Notion of geometric sparsity — accounting for sparse matrix representation of linear maps using underlying geometry of problem.

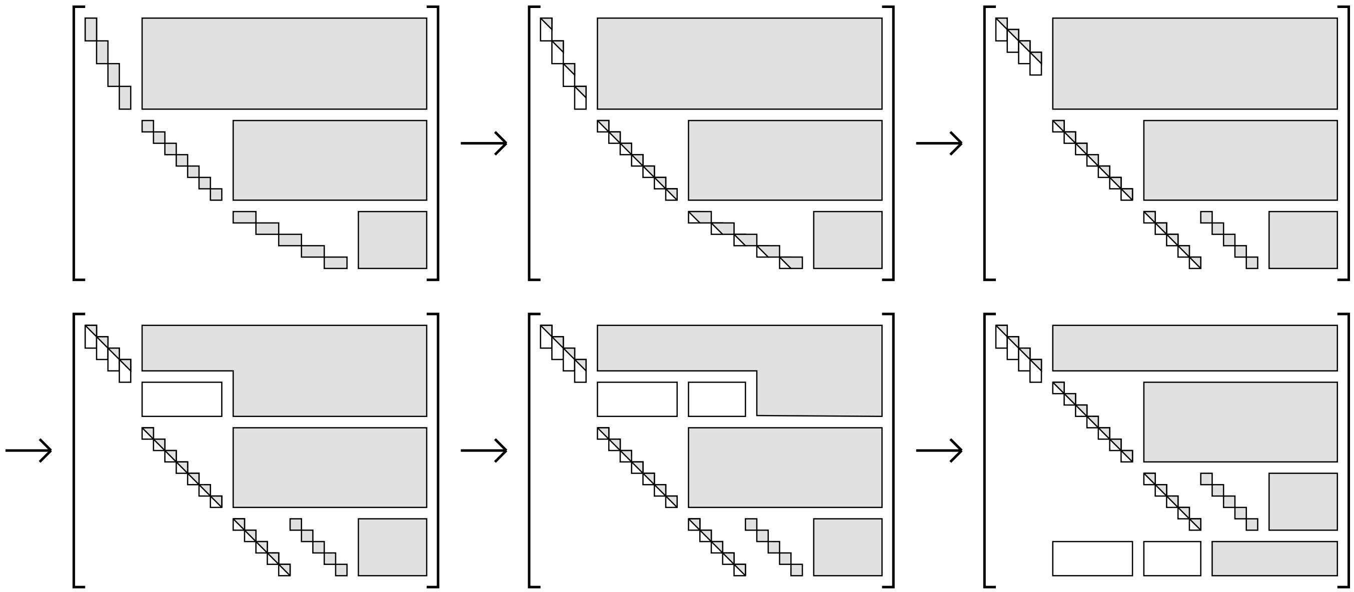
A *geometrically sparse* matrix with a scale parameter r_0 has the following form modulo row- and column-permutations:



Strongly filtered matrix: block triangular where the diagonal blocks are themselves block decomposed.

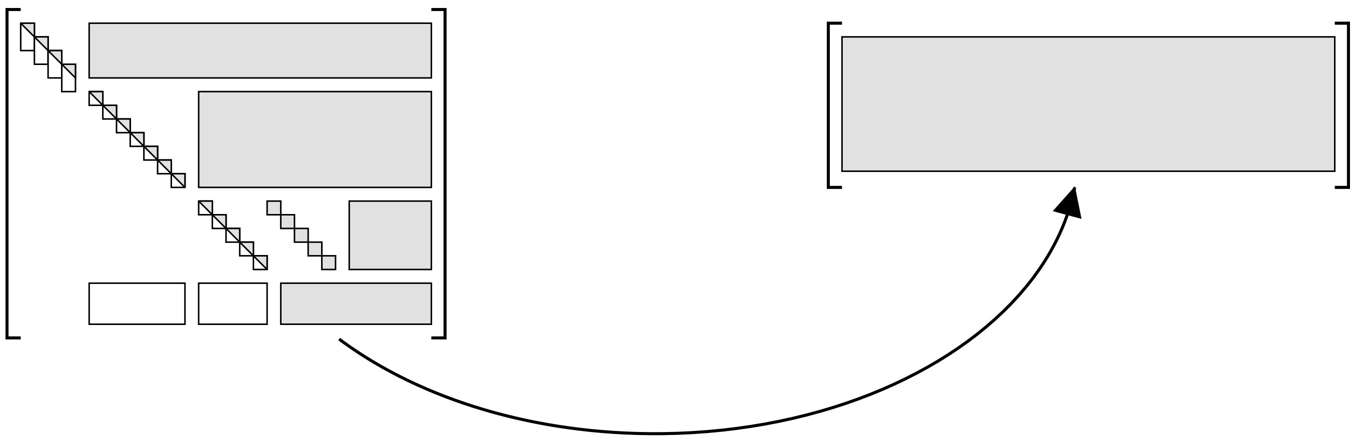
Basic Idea II: Decomposing Strongly Filtered Matrices

The block structure of strongly filtered matrices suggests natural algorithms for LU and QR decomposition.



Basic Idea III: Multiscale Character

The ‘remaining’ submatrix on the bottom right is geometrically sparse with a larger scale parameter $r_1 > r_0$.



The same process may be repeated to this submatrix.

Geometrically Sparse Matrices

'Definition' (Wilkinson). A *sparse matrix* is any matrix with enough zeros that it pays to take advantage of them.

Attempt to give a more concrete definition that accounts for how the sparseness arise.

Definition. A matrix $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ is *geometrically sparse* with scale parameter r if there exist maps $\varphi : \{1, \ldots, m\} \to X$ and $\psi : \{1, \ldots, n\} \to X$ maps sending row and column indices of A into a metric space (X, d) so that $a_{ij} = 0$ whenever $d(\varphi(i), \psi(j)) > r$.

Slogan I

Conjecture. Sparse matrices that arise from physical problems are naturally geometrically sparse or perturbations of geometrically sparse matrices.

The metric space (X, d) is suggested by the problem at hand.

Examples

A banded with bandwidth $2\ell + 1$:
$X = \mathbb{Z}$ or $\mathbb{R}$ with usual metric $|\cdot|$, $a_{ij} = 0$ if $|i - j| > \ell$.

$A = \text{diag}[A_1, \ldots, A_n]$ block diagonal (with square blocks):
$X = \{1, \ldots, n\}$ with discrete metric δ, $a_{ij} = 0$ if $\delta(\varphi(i), \varphi(j)) > 0$.

$$A = \begin{bmatrix} \times & \times & & & \times \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \ddots & & \\ \times & & & \times & \times \end{bmatrix} \in \mathbb{R}^{n \times n}:$$

$X = S^1$ with usual (Riemannian) metric on circle d, $\varphi : \{1, \ldots n\} \to S^1$, $i \mapsto (\cos(2\pi i/n), \sin(2\pi i/n))$, $a_{ij} = 0$ if $d(\varphi(i), \varphi(j)) > 1$.

Examples: Computational Topology

Σ simplicial complex embedded in $\mathbb{R}^n$. To compute $H_k(\Sigma)$, need to find null space of boundary map

$$\partial_k : C_k(\Sigma) \to C_{k-1}(\Sigma),$$
$$[v_0, \ldots, v_k] \mapsto \sum_{i=0}^{k} (-1)^i [v_0, \ldots, \widehat{v}_i, \ldots, v_k].$$

Basis for $C_k(\Sigma)$: k-simplices;
basis for $C_{k-1}(\Sigma)$: $(k-1)$-simplices;
$X = \mathbb{R}^n$;
φ: label each k-simplex by its barycenter;
ψ: label each $(k-1)$-simplex by its barycenter;
$\ell =$ maximal diameter of any simplex of Σ.

Then the matrix representation of ∂_k is geometrically sparse with scale ℓ (likewise for the Laplacian $\Delta := \delta\partial + \partial\delta$).

Examples: Numerical PDE

Finite Difference Methods: discrete approximations of partial differential operators are geometrically sparse.

Finite Element Methods: stiffness matrices in Galerkin's method are geometrically sparse.

Čech complex

X a space, $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ a covering of X.

Nerve of the covering is an abstract simplicial complex:

$$\begin{array}{rcl} \text{vertex} & \longleftrightarrow & U_\alpha \neq \varnothing \\ \text{edge} & \longleftrightarrow & U_\alpha \cap U_\beta \neq \varnothing \\ & \vdots & \\ d\text{-simplex} & \longleftrightarrow & U_{\alpha_0} \cap \cdots \cap U_{\alpha_k} \neq \varnothing \end{array}$$

Also called *Čech complex* and denoted by $\check{\mathsf{C}}(\mathcal{U})$.

$\check{\mathsf{C}}(\mathcal{U})$ is an 'approximation' of X topologically: for nice spaces X, may choose $\mathcal{U}$ so that $\check{\mathsf{C}}(\mathcal{U})$ is homotopy equivalent to X (in particular $H_*(X) \equiv H_*(\check{\mathsf{C}}(\mathcal{U}))$.

Sparse Matrices

‘Definition’ (Wilkinson). A *sparse matrix* is any matrix with enough zeros that it pays to take advantage of them.

Main savings in using sparse matrix algorithms and data structures come in:
time — avoid floating point operations on zero entries;
memory — avoid storing zero entries.

In decomposing a sparse matrix A, e.g. $A = LU$ or $A = QR$, we would like to ensure that the corresponding factors L, U, Q, R are similarly sparse.

LU and QR Decompositions

$A \in \mathbb{R}^{m \times n}$. Want to determine, in a numerically stable fashion, the decompositions

$$P_1 A P_2 = LU \qquad \text{or} \qquad P_1 A P_2 = QR$$

where L lower-triangular; U, R upper-triangular; Q orthogonal; P_1, P_2 permutations.

LU — Gaussian elimination with complete pivoting
QR — similar, with rotations and reflections in place of elementary transformations

Engineering Applications

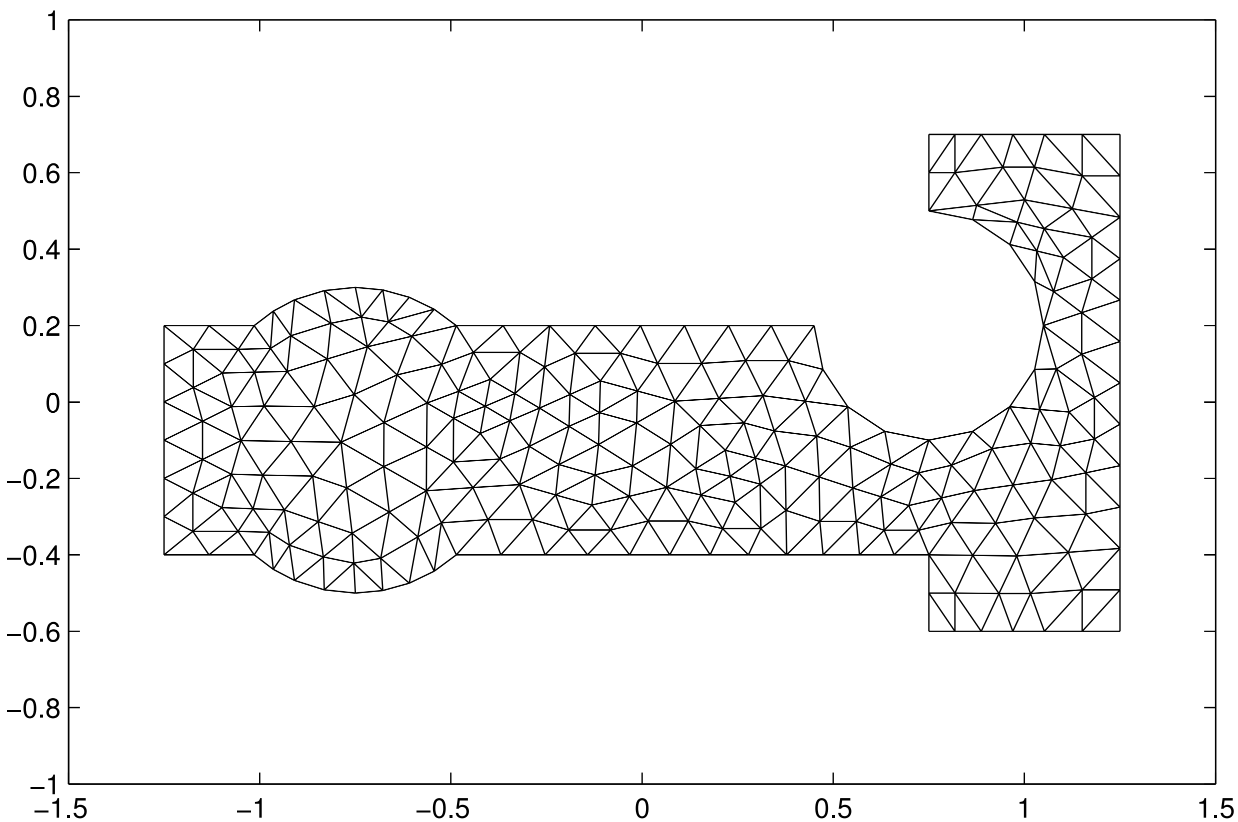
Example. Numerical solutions of partial differential equations typically involve one of the following techniques:

Finite Difference Schemes – approximate the partial differential operators by finite difference operators and then solve the resulting system of equations;
Finite Element Methods – decompose the domain of interest into simpler pieces (e.g. triangulation); approximate the solution by linear combinations of simpler functions supported on these pieces (e.g. splines) and then solve the system of equations coming from the variational formulation of the PDE.

When the PDE is linear, the linear system of equations yields a geometrically sparse matrix.

Using the FEM

We solve $-\nabla \cdot (c\nabla u) + au = f$ with Dirichlet boundary conditions $u = 0$ on the straight edges and Neumann boundary conditions $c\partial u/\partial \nu = -5$ on the circular arcs. Delaunay triangulation algorithm yields:



A Mesh Ordering Strategy