Index of Notation and Definitions

CS 290 H: Graph Laplacians and Spectra

Version of June 12, 2014

There is a lot of variation in terminology and notation in the field of Laplacian matrix computation and spectral graph theory. Indeed, even "Laplacian matrix" is defined differently by different authors!

This list gives the versions of notation, terminology, and definitions that we will use in CS 290H. I mostly follow the conventions of Dan Spielman's 2012 lecture notes, though I prefer not to use greek letters for vectors. I will keep adding to this list during the quarter.

1. Unless otherwise stated, a *graph* $G = (V, E)$ is always an undirected graph whose n vertices are the integers 1 through n, with no multiple edges or loops.

2. The *degree* of a vertex is the number of edges incident on it, or equivalently (because we don't allow multiple edges or loops) the number of its neighboring vertices.

3. A graph is said to be *regular* if every vertex has the same degree.

4. A graph is said to be *connected* if, for every choice of two vertices i and j, there is a *path* of edges from i to j. The *connected components* of a graph are its maximal connected subgraphs.

5. K_n is the *complete graph*, which has n vertices and all $n(n-1)/2$ possible edges.

6. P_n is the *path graph*, which has n vertices and $n-1$ edges in a single path.

7. S_n is the *star graph*, which has n vertices, one with degree $n-1$ and $n-1$ with degree 1.

8. H_k is the *hypercube graph*, which has $n = 2^k$ vertices, all of degree k. Vertices i and j have an edge between them if i and j differ by a power of 2. Equivalently, we can identify each vertex with a subset of $\{1, \ldots, k\}$, with edges to just those subsets formed by adding or deleting one element.

9. G_e or $G_{(i,j)}$ is the graph with n vertices and only one edge $e = (i, j)$.

10. We will write a *vector* as a lower-case latin letter, possibly with a subscript, like x or w_2. We often think of an n-vector as a set of labels for the n vertices of a graph; in that case element i of vector x is written as $x(i)$, and we may write $x \in \mathbb{R}^V$ instead of $x \in \mathbb{R}^n$. In linear algebraic expressions, vectors are column vectors.

11. Two special vectors are $\mathbf{0}$, the vector of all zeros, and $\mathbf{1}$, the vector of all ones.

12. If i is a vertex then $\mathbf{1}_i$ is the *characteristic vector* of i, which is zero except for $\mathbf{1}_i(i) = 1$. Similarly if S is a set of vertices, then $\mathbf{1}_S$ is the vector that is equal to one on the elements of S and zero elsewhere.

13. If d is an n-vector, $\mathrm{diag}(d)$ is the n-by-n diagonal matrix with the elements of d on the diagonal. If A is any n-by-n matrix, $\mathrm{diag}(A)$ is the n-vector of the diagonal elements of A.

14. The *Laplacian* of graph G is the n-by-n matrix L whose diagonal element $L(i,i)$ is the degree of vertex i, and whose off-diagonal element $L(i,j)$ is -1 if $(i,j) \in E$ and 0 if $(i,j) \notin E$. This matrix, which we (and Spielman) just call the Laplacian, is sometimes called the *combinatorial Laplacian* to distinguish it from the normalized Laplacian defined below (44). Note that $L\mathbf{1} = \mathbf{0}$.

15. L_e or $L_{(i,j)}$ is the n-by-n Laplacian matrix of the graph with n vertices and only one edge $e = (i,j)$. This matrix has only four nonzero elements, two 1's on the diagonal and two -1's in positions (i,j) and (j,i); thus

$$L_{(i,j)} = (\mathbf{1}_i - \mathbf{1}_j)(\mathbf{1}_i - \mathbf{1}_j)^T.$$

    The Laplacian of any graph $G = (V,E)$ is the sum of the Laplacians of its edges,

$$L_G = \sum_{e \in E} L_e.$$

16. The *Laplacian quadratic form* (or just LQF) is $x^T L x$, where L is a particular graph's Laplacian and x is a variable n-vector. Its value for a particular vector x is

$$x^T L x = \sum_{(i,j) \in E} (x(i) - x(j))^2.$$

17. A *cut vector* is a vector each of whose elements is $+1$ or -1. We can think of a cut vector x as representing a *cut* that partitions the vertices of graph into two sets $S = \{\, i : x(i) = 1 \,\}$ and $V - S = \{\, i : x(i) = -1 \,\}$. The LQF evaluated at a cut vector is easily seen to be four times the number of edges that cross the cut:

$$x^T L x = 4 \cdot |\{\, (i,j) \in E : i \in S \wedge j \in V - S \,\}|.$$

18. A square matrix Q is *orthogonal* if $Q^T Q = I$, that is, its inverse is its transpose. As vectors, the columns of Q have unit length and are pairwise perpendicular; the same is true of the rows of Q.

19. If $Aw = \lambda w$ for any square matrix A, nonzero vector w, and scalar λ, then λ is an *eigenvalue* of A and w is an *eigenvector* associated with λ.

20. If A is square and B is nonsingular, then the eigenvalues of BAB^{-1} are the same as those of A, and the eigenvectors of BAB^{-1} are B times the eigenvectors of A.

21. If the n-by-n matrix A is symmetric, then it possesses n real eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ (possibly including duplicates) associated with n mutually orthogonal unit-length eigenvectors $w_1, w_2, \ldots, w_n$. If W is the matrix $[w_1\ w_2\ \ldots\ w_n]$ and Λ is the matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ then we can summarize this as $AW = W\Lambda$ and $W^TW = I$. We also have $A = W\Lambda W^T$, whence

$$A = \sum_{i=1}^{n} \lambda_i w_i w_i^T.$$

22. If symmetric A and its eigenvalues and eigenvectors are as in (21), any vector x can be written as a linear combination of eigenvectors,

$$x = \sum_{i=1}^{n} \alpha_i w_i,$$

where $\alpha_i = w_i^T x$. Multiplication by A acts termwise on such a sum:

$$A^k x = \sum_{i=1}^{n} \alpha_i \lambda_i^k w_i.$$

23. If symmetric A and its eigenvalues and eigenvectors are as in (21), the *pseudoinverse* of A is

$$A^\dagger = \sum_{\lambda_i \neq 0} \frac{1}{\lambda_i} w_i w_i^T,$$

where the sum is taken over the nonzero eigenvalues of A. If A is nonsingular, $A^\dagger = A^{-1}$. If x is orthogonal to the null space of A (i.e. $x^T w_i = 0$ whenever $\lambda_i = 0$), then

$$A^\dagger A x = A A^\dagger x = x.$$

24. Every Laplacian L is *positive semidefinite*, which means that none of its eigenvalues are negative. Zero is an eigenvalue of L with multiplicity equal to the number of connected components of the graph G. Therefore, if G is connected, we have $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$. In that case the eigenvector w_1 is the constant vector $\mathbf{1}/\sqrt{n}$.

25. The *Fiedler value* of a graph is λ_2, its second-smallest eigenvalue, and the *Fiedler vector* is w_2, the associated eigenvector. The Fiedler value of a graph is also called its *algebraic connectivity*. Note that $\lambda_2 = 0$ iff the graph is not connected.

26. The *positive semidefinite square root* of a positive semidefinite matrix A with eigenvalues and eigenvectors as in (21) is the matrix

$$A^{1/2} = \sum_{i=1}^{n} \lambda_i^{1/2} w_i w_i^T.$$

We write the psd square root of $A^\dagger$ as

$$A^{\dagger/2} = \sum_{\lambda_i \neq 0} \lambda_i^{-1/2} w_i w_i^T.$$

27. The *trace* of a matrix is the sum of its diagonal elements. The trace is also equal to the sum $\sum_i \lambda_i$ of its eigenvalues. The trace of a Laplacian is equal to twice the number of edges in the graph.

28. **Gershgorin's theorem**. If A is any square matrix (real or complex), its n eigenvalues are all contained in the union of the n disks $D_1, \ldots, D_n$ in the complex plane defined by

$$D_i = \{\alpha : |\alpha - A(i,i)| \leq \sum_{j \neq i} |A(i,j)|\}.$$

This implies, for example, that the largest eigenvalue λ_n of a Laplacian is at most twice the maximum vertex degree.

29. The *Rayleigh quotient* of a nonzero vector x and a matrix A is

$$\frac{x^T A x}{x^T x}.$$

If $Ax = \lambda x$, then the Rayleigh quotient of x and A is λ.

30. **Rayleigh quotient theorem**. The eigenvectors of a symmetric matrix A are critical points of its Rayleigh quotient (considered as a real-valued function of an n-vector). Specifically,

$$\lambda_k = \min_{x \perp w_1, \ldots, w_{k-1}} \frac{x^T A x}{x^T x} = \max_{x \perp w_{k+1}, \ldots, w_n} \frac{x^T A x}{x^T x},$$

and the extreme values are attained at $x = w_k$. In particular, therefore, for a Laplacian L the Fiedler value is

$$\lambda_2 = \min_{x \perp \mathbf{1}} \frac{x^T A x}{x^T x},$$

attained at the Fiedler vector w_2.

31. **Courant-Fischer theorem** (a stronger version of the Rayleigh quotient theorem). The eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$ of a symmetric matrix A are characterized by

$$\lambda_k = \max_{\dim \mathbb{S} = n-k+1} \min_{x \in \mathbb{S}} \frac{x^T A x}{x^T x} = \min_{\dim \mathbb{S} = k} \max_{x \in \mathbb{S}} \frac{x^T A x}{x^T x},$$

where $\mathbb{S}$ ranges over subspaces of $\mathbb{R}^n$. The extreme values are attained at $x = w_k$.

32. A *test vector for* λ_2 is an n-vector that is orthogonal to $\mathbf{1}$. By the Raleigh quotient theorem, if v is any test vector then $\lambda_2 \leq v^T L v / v^T v$. Note that any vector x can be converted to a test vector $v = x - (x^T \mathbf{1}/n)\mathbf{1}$; in words, subtracting off the mean of any vector orthogonalizes it against the constant vector.

33. A *weighted graph* is an undirected graph that comes with *positive* weights on the edges, which we write $c(e)$ or $c(i,j)$. Note that $c(i,j) = c(j,i)$.

34. The *weighted Laplacian* of a weighted graph is the n-by-n matrix L whose off-diagonal element $L(i,j)$ is $-c(i,j)$ if $(i,j) \in E$ and 0 if $(i,j) \notin E$, and whose diagonal element $L(i,i) = \sum_{k \neq i} c(i,k)$ is chosen to make the row sums zero. Like the ordinary Laplacian, we have $L\mathbf{1} = \mathbf{0}$, and indeed 0 is an eigenvalue of L with multiplicity equal to the number of connected components of G.

35. **Multiple of a graph**. If G is a graph or a weighted graph and $\alpha > 0$ is a constant, αG is the graph whose edge weights are all multiplied by α; the Laplacian of αG is α times the Laplacian of G.

36. **Semidefinite ordering**. If A is a matrix, $A \succeq 0$ means that A is positive semidefinite. Thus $L \succeq 0$ for any Laplacian L. If A and B are matrices, $A \succeq B$ means $A - B \succeq 0$. If G and H are graphs or weighted graphs, $G \succeq H$ means $L_G \succeq L_H$. Note that $G \succeq H$ iff $x^T L_G x \geq x^T L_H x$ for all vectors x. For matrices $A \succeq 0$ and $B \succeq 0$, $A \succeq B$ implies $\lambda_k(A) \geq \lambda_k(B)$ for all k. Also, $A \succeq B$ implies $B^\dagger \succeq A^\dagger$.

37. **Graph approximation**. For any constant $\alpha \geq 1$, (weighted) graph H is an *α-approximation* of (weighted) graph G if $\alpha H \succeq G \succeq H/\alpha$.

38. The *boundary* of a set $S \subseteq V$ of vertices, written ∂S, is the set of edges with just one endpoint in S. Formally, $\partial S = \{\, (i,j) \in E : i \in S \wedge j \in V - S \,\}$. The number of edges in ∂S is $|\partial S|$.

39. The *isoperimetric number* of a set $S \subseteq V$ of vertices, written $\theta(S)$, is the ratio

$$\phi(S) = \frac{|\partial S|}{\min(|S|, |V - S|)}.$$

    This is one sort of "surface-to-volume ratio"; see the definition of conductance (42) for another.

40. The *isoperimetric number* of a graph G, written θ_G, is $\min_{S \subset V} \theta(S)$, the smallest isoperimetric number of any nonempty proper subset of vertices. Note that $\theta_G = 0$ if and only if G is not connected.

41. **Isoperimetric theorem**. For any set S of vertices, $\theta(S) \geq \lambda_2(1 - |S|/|V|)$. This says that the larger λ_2 is, the larger the surface-to-volume ratio of any relatively small set of vertices must be.

42. The *conductance* of a set $S \subseteq V$ of vertices, written $\phi(S)$, is the ratio

$$\phi(S) = \frac{|\partial S|}{\min(d(S), d(V-S))},$$

where $d(S)$ is the sum of the degrees of the vertices in S. This is another sort of "surface-to-volume ratio"; isoperimetric number (40) measures volume just by counting vertices, while conductance measures volume by counting vertices weighted by their degrees. ("Conductance" has a different meaning in resistive networks; see (56) below.)

43. The *conductance* of a graph G, written ϕ_G, is $\min_{S \subset V} \phi(S)$, the smallest conductance of any nonempty proper subset of vertices. This is sometimes called the "Cheeger constant" of the graph, but definitions are particularly variable here and we'll stick to this one. Note that $\phi_G = 0$ iff G is not connected. ("Conductance" has a different meaning in resistive networks; see (56) below.)

44. The *normalized Laplacian* of graph G is the n-by-n matrix N whose diagonal element $N(i,i)$ is equal to 1, and whose off-diagonal element $N(i,j)$ is $-\sqrt{d(i)d(j)}$, the geometric mean of the degrees of vertices i and j, where we define d to be the vector of vertex degrees of G. Another way to say it is that the normalized Laplacian is the (ordinary) Laplacian with rows and columns scaled symmetrically to make the diagonal elements equal to 1. If $D = \mathrm{diag}(d)$ is the diagonal matrix of degrees, then

$$N = D^{-1/2} L D^{-1/2}.$$

Some authors, including notably Fan Chung in her wonderful book *Spectral Graph Theory*, use the name "Laplacian" for this matrix N instead of for our L.

45. The normalized Laplacian N is symmetric and positive semidefinite, and like the Laplacian it has 0 as an eigenvalue with multiplicity equal to the number of connected components of G. In general however N's eigenvalues and eigenvectors are different from L's. We write $0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_n$ for the eigenvalues of N. The eigenvector corresponding to ν_1 is not the constant vector, but the vector $d^{1/2}$ of the square roots of the vertex degrees:

$$N d^{1/2} = D^{-1/2} L D^{-1/2} d^{1/2} = D^{-1/2} L \mathbf{1} = D^{-1/2} \mathbf{0} = \mathbf{0}.$$

46. It follows from Gershgorin's theorem (28) that the eigenvalues of the normalized Laplacian N are always bounded by 0 and 2,

$$0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_n \leq 2.$$

47. The analysis in (45) above shows, incidentally, that the normalized Laplacian is *not* a weighted Laplacian according to definition (34) unless the graph is regular. There is a way to define a "normalized weighted Laplacian" that we may use later on.

48. **Cheeger's inequality**. The normalized Laplacian can be used to give both upper and lower bounds on the conductance,

$$\nu_2/2 \leq \phi_G \leq \sqrt{2\nu_2}.$$

Equivalently,

$$\phi_G^2/2 \leq \nu_2 \leq 2\phi_G.$$

The upper bound on ν_2 is analogous to the isoperimetric inequality (41). The lower bound on ν_2 is Cheeger's inequality, one of the most significant theorems of spectral graph theory.

49. **Cauchy-Schwarz inequality**. Just for reference, because it comes up in several of the proofs we're looking at. If x and y are n-vectors, then

$$|x^T y| \leq \|x\| \, \|y\|.$$

Equivalently,

$$\Big(\sum_i x(i)y(i)\Big)^2 \leq \Big(\sum_i x(i)^2\Big)\Big(\sum_i y(i)^2\Big).$$

50. The k-dimensional *Krylov subspace* based on a square matrix A and a vector b is

$$\mathcal{K}_k(A, b) = \text{span}(b, Ab, A^2b, \ldots, A^{k-1}b).$$

51. Let A be n-by-n and symmetric, and let $k \leq n$. Let Q_k be n-by-k with orthonormal columns (i.e. $Q_k^T Q_k = I$) such that $Q_k^T A Q_k = T_k$ is symmetric and tridiagonal. Let $T_k = V_k \Theta V_k^T$ where V_k is k-by-k and orthogonal, and $\Theta = \text{diag}(\theta_1, \ldots, \theta_k)$ is diagonal. Then $\theta_1, \theta_2 \ldots, \theta_k$ (the eigenvalues of T_k) are *Ritz values* for A, and the columns of $Q_k V_k$ are the corresponding *Ritz vectors*.

52. If matrix M is partitioned into 2-by-2 block form

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

such that A is square and nonsingular, then the *Schur complement of A in M* is the matrix

$$S = D - CA^{-1}B.$$

53. If G is a graph with n vertices and m edges, an *incidence matrix* of G is an n-by-m matrix U with a column for each edge of G. The column for edge (i, j) contains two nonzeros, a 1 and a -1, one in row i and one in row j. The incidence matrix is not unique; permuting its columns or negating some of its columns produces another incidence matrix. Any incidence matrix U is related to the Laplacian L by

$$L = UU^T.$$

54. If G is a weighted graph, its incidence matrix is the same as above (53), without weights. If $C = \mathrm{diag}(c)$ is the diagonal matrix of edge weights, in the same order as the columns of U, then the weighted Laplacian L satisfies

$$L = UCU^T.$$

55. An *augmented matrix* of an unweighted graph with n vertices and m edges is a symmetric $(n+m)$-by-$(n+m)$ matrix defined in block form as

$$\begin{pmatrix} I & U^T \\ U & 0 \end{pmatrix},$$

where U is an incidence matrix and I is the m-by-m identity matrix. An augmented matrix of a weighted graph is

$$\begin{pmatrix} R & U^T \\ U & 0 \end{pmatrix},$$

where $R = \mathrm{diag}(1/c) = C^{-1}$ is the diagonal matrix of inverse edge weights. The Schur complement of R is then $-UR^{-1}U^T = -L$, the negative of the weighted Laplacian.

56. A *resistive network* is a weighted graph with n vertices interpreted as nodes of an electrical circuit and m edges interpreted as resistors joining pairs of nodes. If the resistor at edge e has resistance $r(e)$, the edge's weight is the inverse resistance $c(e) = 1/r(e)$. (Inverse resistance is often called "conductance," whence the letter c, but we will not use the term in this context to avoid confusion with the unrelated notion of graph conductance in (42) above.)

57. In a resistive network G, suppose a current $b(i)$ is injected at each node i, where $b^T\mathbf{1} = 0$ so the same total current is injected and removed from the network as a whole. For each edge $(i,j) \in E$ with $i < j$, let $f(i,j)$ be the *current* or *flow* along edge (i,j). Then *Ohm's law*, or current times resistance equals voltage, says

$$v(i) - v(j) = f(i,j)r(i,j) \quad \text{for all } (i,j) \in E.$$

Kirchoff's current law, or current entering a node equals current leaving the node, says

$$\sum_{j:(i,j)\in E} f(i,j) = b(i) \quad \text{for all } i \in V.$$

The two laws can be combined in one augmented system (55) as

$$\begin{pmatrix} R & U^T \\ U & 0 \end{pmatrix} \begin{pmatrix} f \\ -v \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix},$$

where U is the incidence matrix of G with appropriate signs, and R is the m-by-m diagonal matrix of edge resistances. The Schur complement of R is $-UR^{-1}U^T$, which is the negative

Laplacian $-L$. A step of block Gaussian elimination on the augmented system thus leads to the Laplacian linear system

$$Lv = b$$

relating the node voltages to the externally injected currents.

58. In a resistive network, the *effective resistance* between two vertices i and j, written $R^{\text{eff}}(i,j)$, is the positive difference in voltage between i and j when one unit of current is injected at i and extracted at j. That is, if $u_{ij} = \mathbf{1}_i - \mathbf{1}_j$ for $i < j$ is the vector whose k'th element is equal to 1 when $k = i$, equal to -1 when $k = j$, and equal to zero elsewhere, and $Lv = u_{ij}$, then

$$R^{\text{eff}}(i,j) = u_{ij}^T v = u_{ij}^T L^\dagger u_{ij} = v^T L v.$$

We write $R_G^{\text{eff}}(i,j)$ if the graph is not clear from context.

59. Let G and H be resistive networks on the same number of vertices. Then

$$R_G^{\text{eff}}(i,j) \geq R_H^{\text{eff}}(i,j) \text{ for all } i,j \in V \quad \text{if and only if} \quad H \succeq G.$$

60. The *characteristic polynomial* of n-by-n matrix A is the degree-n polynomial in one variable λ defined by

$$P(\lambda) = \det(\lambda I - A).$$

As suggested by the notation, the n roots of $P(\lambda) = 0$ are the n eigenvalues $\lambda_1, \ldots, \lambda_n$ of A.

61. **Cayley-Hamilton theorem**. For any matrix A, the characteristic polynomial interpreted as a matrix polynomial and evaluated at A gives the zero matrix; that is, $P(A) = 0$.

62. The *condition number* of a square matrix A is $\kappa(A) = \|A\| \, \|A^{-1}\|$, interpreted as ∞ if A is singular. If A is symmetric and positive definite, $\kappa(A) = \lambda_n/\lambda_1$ is the ratio of the extreme eigenvalues.

63. **Conjugate gradient**. The conjugate gradient algorithm (or CG) solves $Ax = b$, where $A \succ 0$ is a symmetric, positive definite matrix (see Shewchuk for details). Each iteration performs one matrix-vector multiplication with A and some vector arithmetic, taking $O(n+m)$ time per iteration if A has m nonzeros. The relative error in the approximate solution x_j is bounded by

$$\frac{||x_j - x||_A}{||x||_A} < \epsilon$$

after

$$j = O(\sqrt{\kappa(A)} \log(1/\epsilon))$$

iterations (in exact arithmetic), where $\kappa(A) = \lambda_n/\lambda_1$ is the condition number of A and $||v||_A = (v^T A v)^{1/2}$ is the A-norm. With some care, CG can also be used for a positive semidefinite matrix whose null space is known, e.g. a weighted graph Laplacian.

64. **Preconditioned conjugate gradient**. The preconditioned conjugate gradient algorithm (or PCG) solves $Ax = b$ by applying CG to the linear system

$$(B^{-1/2}AB^{-1/2})(B^{1/2}x) = B^{-1/2}b,$$

where A and B are symmetric positive definite. Each iteration of PCG performs one matrix-vector multiplication with A, one linear system solve with B, and some vector arithmetic. Matrix B is called a *preconditioner* for A, and may or may not be formed explicitly. A good preconditioner satisfies two criteria:

- It should be "easy" to solve the linear system $By = z$ for y.
- The condition number $\kappa(B^{-1/2}AB^{-1/2}) = \kappa(AB^{-1})$ should be smaller than $\kappa(A)$.

With some care, PCG can also be used with positive semidefinite matrices A and B if they have the same null space.

65. For a symmetric positive semidefinite matrix A, the *finite condition number* is $\kappa_f(A) = \lambda_n/\lambda_k$, where λ_k is the smallest nonzero eigenvalue. For example, if L is the Laplacian of a connected graph, $\kappa_f(L) = \lambda_n/\lambda_2$ is the relevant condition number for the convergence of conjugate gradient.

66. Let A and B be symmetric positive semidefinite matrices with the same null space (e.g., weighted Laplacians of two connected graphs on the same vertices). The *finite condition number* $\kappa_f(A, B)$ is $\kappa_f(AB^\dagger)$, which is the relevant condition number for the convergence of conjugate gradient on $Ax = b$ with preconditioner B. Note that $\kappa_f(A, B) = \kappa_f(B, A)$.

67. Let A and B be two symmetric positive semidefinite matrices with the same null space. If $\alpha B \preceq A \preceq \beta B$, then $\alpha \leq \lambda \leq \beta$ for every nonzero eigenvalue λ of $AB^\dagger$, and therefore $\kappa_f(A, B) \leq \beta/\alpha$.

68. **Symmetric Gaussian elimination**. If A is a positive definite matrix (or, with some care, a positive semidefinite matrix), the *Cholesky factorization* is $A = R^T R$, where R is an upper triangular matrix with positive diagonal elements (non-negative in the semidefinite case). The Cholesky factorization of any n-by-n matrix can be computed in $O(n^3)$ time and $O(n^2)$ memory; some but by no means all sparse matrices have better bounds.

69. **Cholesky graph game**. Given positive (semi)definite A with undirected graph $G(A)$, the undirected graph $G^+(A) = G(R + R^T)$ of the Cholesky factors of A is obtained as follows:

```
for j = 1 : n
  mark vertex j;
  add "fill" edges between the umarked neighbors of vertex j;
end for
```

(This gives the nonzero structure of R but not the nonzero values.) We are free to mark the vertices in any order; choosing a different order corresponds to applying a permutation symmetrically to the rows and columns of A.

70. **Parter's theorem**. If the graph $G(A)$ is a tree, a vertex ordering exists for which the Cholesky factorization adds no fill and solving $Ax = b$ takes only $O(n)$ time and memory.

71. **Nested dissection**. If the graph $G(A)$ is the $\sqrt{n}$-by-$\sqrt{n}$ grid graph, the best possible elimination ordering has $O(n \log n)$ fill, for which Cholesky takes $O(n^{3/2})$ time. The same upper bounds hold for any planar graph. For the three-dimensional grid graph, the best possible fill is $O(n^{4/3})$ and Cholesky takes $O(n^2)$ time.

72. **Spanning trees, fundamental cycles**. If A is a connected graph (or a connected weighted Laplacian), a *spanning tree* of A is a subgraph B of A that has no cycles (a tree) and includes all the vertices of A (spanning). Edges of B have the same weight in B as they have in A. Each edge $e = (i, j)$ that is in A but not in B induces a *fundamental cycle* consisting of e and the unique path $P(e)$ in B between its endpoints i and j. For an edge $e = (i, j)$ of B, we write $P(e) = (i, j)$ for the length-one path between its endpoints, but e does not induce a fundamental cycle.

73. **Dilation**. If B is a spanning tree of A (connected weighted Laplacians), the *dilation* of an edge e of A is the number of edges on the path $P(e)$ between e's endpoints in B. We define

$$\text{dilation}(A, B) = \max_{e \in E(A)} \text{dilation}(e),$$

the largest dilation of any edge of A, and informally call this "the dilation of B."

74. **Congestion**. If B is a spanning tree of A (connected weighted Laplacians), the *congestion* of an edge f of B is the number of edges e of A (including f) for which f is on the path $P(e)$ between e's endpoints in B. We define

$$\text{congestion}(A, B) = \max_{f \in E(B)} \text{congestion}(f),$$

the largest congestion of any edge of B, and informally call this "the congestion of B."

75. **Stretch**. If B is a spanning tree of A (connected Laplacians with edge weights $c(e)$), the *stretch* of an edge e of A is the weighted version of its dilation:

$$\text{stretch}(e) = c(e) \sum_{f \in P(e)} \frac{1}{c(f)}.$$

If we interpret A and B as resistive networks with resistances $r(e) = 1/c(e)$, then the stretch of edge e is just $r(P(e))/r(e)$.

76. **Maximum-weight spanning trees**. The *weight* of a spanning tree B of A (connected Laplacians with edge weights $c(e)$), is the total weight of the edges of B,

$$\sum_{e \in E(B)} c(e).$$

A *maximum-weight spanning tree* of A is a spanning tree with the largest possible weight. Such a tree can be found by a simple greedy algorithm in $O(m \log n)$ time (or a little faster by a more complicated algorithm) if A has m edges.

77. **Vaidya's tree theorem**. If B is a maximum-weight spanning tree of A (connected weighted Laplacians), then the finite condition number of the preconditioned system satisfies

$$\kappa_f(A, B) \leq \text{dilation}(A, B) \cdot \text{congestion}(A, B).$$

This is at most nm if graph A has n vertices and m edges.

78. **Low-stretch spanning trees**. The *stretch* of a spanning tree B of A (connected weighted Laplacians) is the total stretch of the edges of A (not the edges of B),

$$\text{stretch}(A, B) = \sum_{e \in E(A)} \text{stretch}(e).$$

A nontrivial theorem (see Spielman's 2010 survey paper in the course references) is that every weighted graph A with n vertices and m edges has a spanning tree B with nearly linear stretch,

$$\text{stretch}(A, B) = O(m \log n \log \log n (\log \log \log n)^3).$$

It is an open problem whether this can be improved to $O(m \log n)$ in general, which would be best possible. The tree B can be computed in time $O(m \log n + n \log^2 n)$.