* Method of Multipliers (Augmented Lagrangian)

$$\min \; -\frac{1}{2}(x_1^2 + x_2^2)$$

$$\text{s.t.} \quad x_1 + 2x_2 = 3$$

$$L_c(x,\nu) = -\frac{1}{2}(x_1^2+x_2^2) + \nu(x_1+2x_2-3) + \frac{c}{2}(x_1+2x_2-3)^2 \;, \quad c \geq 1$$

Nice regularization of optimization landscape → can convert non-convex problems to potentially sc problems.

At iteration k:

$$x_k = \arg\min L_{c_k}(x,\nu_k)$$
$$\nu_{k+1} = \nu_k + c_k h(x_k)$$
$$c_{k+1} = \beta c_k \quad , \beta > 1$$

* Dual-ascent:

$$\min f(x)$$
$$\text{s.t.} \quad Ax = b$$

At iteration k:

$$x_k = \arg\min_x f(x) + \nu^T Ax$$
$$\nabla g(\nu_k) = \underbrace{Ax_k - b}_{\text{residual}}$$
$$\nu_{k+1} = \nu_k + \alpha_k \nabla g(\nu_k)$$

* Dual Decomposition:

$$\min_{\{x_i \in \mathbb{R}^{n_i}\}} \sum_{i=1}^{B} f_i(x_i)$$

$$\text{s.t.} \quad Ax = b$$

$$A = [A_1 \; A_2 \; \ldots \; A_B] \qquad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_B \end{bmatrix}$$

(Broadcast)

Scatter-Gather

$$\rightarrow x_i^{(k)} = \arg\min f_i(x) + \nu_k^T A_i x \quad \forall i$$

$$\nu_{k+1} = \nu_k + \alpha_k \left(\sum_{i=1}^{B} A_i x_i^{(k)} - b \right)$$

Parallelizability

$$\left.\begin{array}{ll} \min & f(x) \\ \text{s.t.} & Ax = b \end{array}\right\}$$

Dual ascent with "augmented Lagrangian"

$$L_c(x, \nu) = f(x) + \nu^T(Ax - b) + \frac{c}{2}\|Ax - b\|^2$$

* Can we enjoy the best of both worlds? ⟶ Yes

Augment Lagrangian's optimization landscape

Parallelizable (Dual-ascent)

ADMM (Alternating direction method of Multipliers)

* Consider the problem of the form:

$$\min_{x,z} \; f(x) + g(z)$$

$$\text{s.t.} \quad Ax + Bz = c$$

$$L_\rho(x, z, \nu) = f(x) + g(z) + \nu^T(Ax + Bz - c) + \frac{\rho}{2}\|Ax + Bz - c\|^2$$

At iteration k: (Using Method of Multipliers)

$$\left.\begin{array}{l} (x_k, z_k) = \arg\min_{(x,z)} L_\rho(x, z, \nu_k) \\ \nu_{k+1} = \nu_k + \rho(Ax_k + Bz_k - c) \end{array}\right\}$$

ADMM repeats the step: $(x^{(k-1)}, z^{(k-1)}, \nu_k)$

$$\boxed{\begin{array}{l} x^{(k)} = \arg\min_x L_\rho(x, z^{(k-1)}, \nu_k) \\ z^{(k)} = \arg\min_z L_\rho(x^{(k)}, z, \nu_k) \\ \nu_{k+1} = \nu_k + \rho(Ax^{(k)} + Bz^{(k)} - c) \end{array}}$$

* Convergence guarantees:

Under modest assumptions on f, g (do not require A, B to be full rank),

then ADMM iterates satisfy for any $\rho > 0$,

↳ Residual convergence:

$$r^{(k)} = Ax^{(k)} + Bz^{(k)} - c \longrightarrow 0 \quad \text{as } k \to \infty$$

↳ Objective convergence:

$$f(x^{(k)}) + g(z^{(k)}) \longrightarrow f^* + g^* \quad \text{as } k \to \infty$$

↳ Dual convergence:

$$\nu^{(k)} \longrightarrow \nu^* \quad \text{as } k \to \infty$$

* Scaled-form ADMM: $\quad w := \nu/\rho$

$$L_\rho(x, z, \nu) = f(x) + g(z) + \nu^T(Ax + Bz - c) + \frac{\rho}{2}\|Ax + Bz - c\|^2$$

$$= f(x) + g(z) + \frac{2\rho}{2}\left(\frac{\nu}{\rho}\right)^T(Ax + Bz - c) + \frac{\rho}{2}\|Ax + Bz - c\|^2$$

$$= f(x) + g(z) + 2\cdot\rho\cdot\frac{w^T}{2}(Ax + Bz - c) + \frac{\rho}{2}\|Ax + Bz - c\|^2 + \frac{\rho}{2}\|w\|^2 - \frac{\rho}{2}\|w\|^2$$

$$= f(x) + g(z) + \frac{\rho}{2}\Big[\|w\|^2 + \|Ax + Bz - c\|^2 + 2w^T(Ax + Bz - c)\Big] - \frac{\rho}{2}\|w\|^2$$

$$= f(x) + g(z) + \frac{\rho}{2}\|Ax + Bz - c + w\|^2 - \frac{\rho}{2}\|w\|^2$$

Final step is:

$$w_{k+1} = w_k + r^{(k)}$$

$$w_{k+1} = w_0 + \text{Sum of residuals.}$$

—— x —— x —— x —— x —— x —— x ——

* Basics of Graph Theory:

convex / SC / PL-inequality

$$\min_{x\in\mathbb{R}^n} f(x)$$

$$x^*$$

$$\sum_{i=1}^{n} f_i(x)$$

$$\min_{\{x_1,\dots,x_N\}\in\mathbb{R}^{nN}} \sum_{i=1}^{N} f_i(x_i)$$

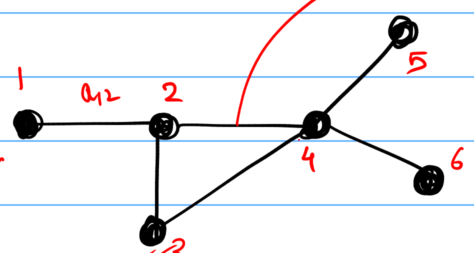
— s.t. $x_1 = x_2 = \dots = x_N$

$$x^*$$

f_i's are private objective functions.

$$f_i := \sum_{(x,y)\sim D_i} \ell(\hat{y}_i, y_i)$$

edges



Undirected, unweighted graph

nodes/ vertices/ agents

$2 \in \mathcal{N}_1$

$\mathcal{N}_i$: neighborhood set of agent i

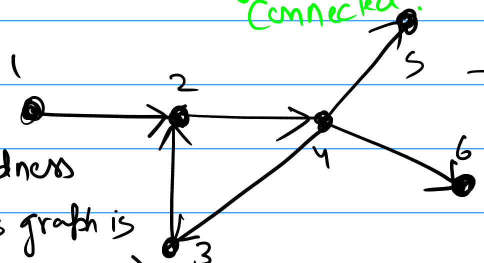
$$a_{ij} = \begin{cases} 1 & \text{if } \exists \text{ an edge b/w } i \text{ and } j \\ 0, & \text{else} \end{cases}$$

$$\mathcal{N}_2 = \{1, 3, 4\}$$

$\mathcal{V}$:= set of vertices $\{1,2,3,4,5,6\}$

$\mathcal{E}$:= set of edges.

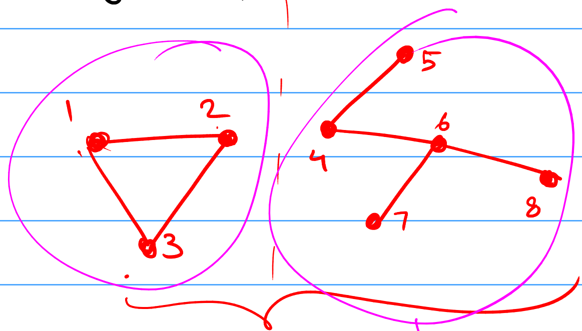
Is this graph connected?



→ Directed graphs (Digraphs) → Directed, weighted graphs

Strong connectedness (This graph is not strongly connected)

*

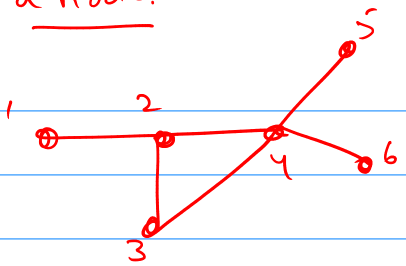


Graph is not connected:

Connected: A graph is connected if there exists a path b/w any pair of nodes.

Subgraphs → Connected components of a graph.

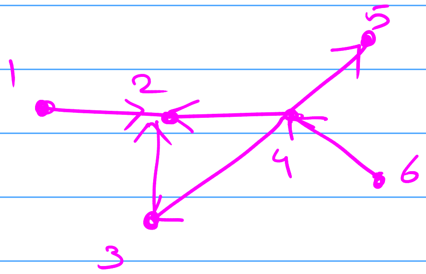
* **Degree of a node:**



$d(1) = 1$
$d(2) = 3$
$d(3) = 2$
$d(4) = 4$
$d(5) = 1$
$d(6) = 1$

$d_i :=$ size of the neighborhood set $\mathcal{N}_i$

Degree matrix of a graph: $D = \begin{bmatrix} 1 & & & & & 0 \\ & 3 & & & & \\ & & 2 & & & \\ & & & 4 & & \\ & & & & 1 & \\ 0 & & & & & 1 \end{bmatrix}$



In-degree $\quad d_{in}(1) = 0$
$\qquad\qquad\;\; d_{out}(1) = 1$

* **Edge Adjacency matrix:** $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$

For each i, $\sum_{j=1}^{N} a_{ij} = d_i \leftarrow$ degree of node i

* **Laplacian:** $L := D - A \}$ symmetric

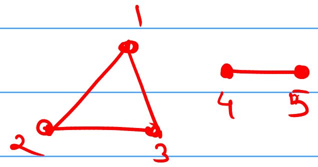
$L_{ij} = d_{ij} - a_{ij} \qquad \sum_{j=1}^{N} l_{ij} = \underbrace{\sum_{j=1}^{N} d_{ij}}_{d_i} - \underbrace{\sum_{j=1}^{N} a_{ij}}_{d_i} = 0$

$L \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = 0 \longrightarrow$ 0 is an eigenvalue of L with $\mathbb{1}_N$ being the corresponding eigenvector.

Normalized Laplacian $\begin{cases} L_{sym} = I - D^{-1/2} A D^{-1/2} \longrightarrow \text{nCut} \\ L_{rw} = I - D^{-1} A \longrightarrow \text{Ratio Cut} \end{cases}$

Spectral Clustering $\leftarrow$

$L\,\mathbb{1} = 0 \longrightarrow$ Is 0 a simple eigenvalue?



Disconnected graph
\# connected components is 2.

$$A = \left[\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right]$$

$$D = \begin{bmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}$$

$$L = D - A = \left[\begin{array}{c|c} L_1 & 0 \\ \hline 0 & L_2 \end{array}\right]$$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Algebraic multiplicity of 0 eigenvalues of L
= \# of connected components.

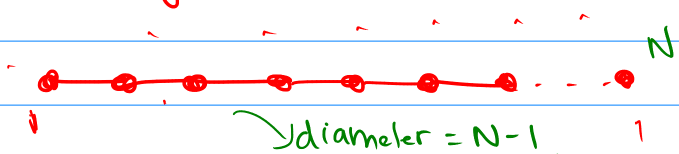
Eigenvalues of Laplacian:

$$0 = \lambda_1 \le \lambda_2 \le \lambda_3 \le \dots$$

Smallest non-zero eigenvalue

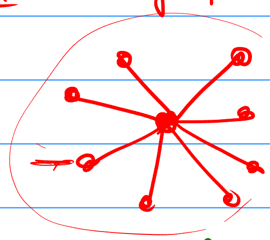
$\lambda_2(L) \rightarrow$ Fiedler eigenvalue.

Ex: Line graph:



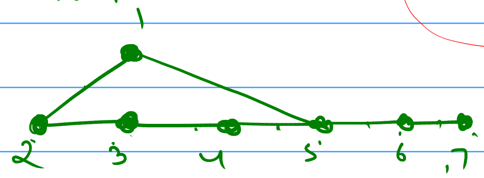
diameter $= N-1$

Ex: star graph:



diameter $= 2$

Diameter of a graph:



$1 \to 2 \quad = 1$

$1 \to 3 \quad = 2$

diameter $= 4$

Brendan McKay   Fiedler value $\ge \dfrac{4}{n\,(\text{diameter})}$ → for connected graphs

* L is a positive semi-definite matrix $\Rightarrow \lambda(L) \geq 0$.

$$x^T L x \geq 0$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} x_i l_{ij} x_j$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} x_i (D_{ij} - A_{ij}) x_j$$

$$= \sum_{i,j=1}^{N} x_i D_{ij} x_j - \sum_{i,j=1}^{N} x_i A_{ij} x_j$$

$$= \sum_{i=1}^{N} x_i^2 d_i - \sum_{i,j=1}^{N} x_i x_j A_{ij}$$

$$= \frac{1}{2} \sum_{i=1}^{N} x_i^2 d_i + \frac{1}{2} \sum_{j=1}^{N} x_j^2 d_j - \sum_{i,j=1}^{N} x_i x_j A_{ij}$$

$$= \frac{1}{2} \sum_{i,j=1}^{N} A_{ij} (x_i - x_j)^2 \geq 0$$