Data Analytics and Machine Learning Group
Department of Informatics
Technical University of Munich

Esolution
Place student sticker here

Note:

- During the attendance check a sticker containing a unique code will be put on this exam.
- This code contains a unique number that associates this exam with your registration number.
- This number is printed both next to the code and to the signature field in the attendance check list.

Machine Learning for Graphs and Sequential Data

Graded Exercise:	IN2323 / Endterm	**Date:**	Friday 30th July, 2021
Examiner:	Prof. Dr. Stephan Günnemann	**Time:**	11:30 – 12:45

Working instructions

- This graded exercise consists of **26 pages** with a total of **19 problems**.
  Please make sure now that you received a complete copy of the answer sheet.
- The total amount of achievable credits in this graded exercise is 140 credits.
- Allowed resources:
  - all materials that you will use on your own (lecture slides, calculator etc.)
  - **not allowed are any forms of collaboration between examinees and plagiarism**
- You have to sign the code of conduct. (Typing your name is fine)
- You have to either print this document and scan your solutions or paste scans/pictures of your handwritten solutions into the solution boxes in this PDF. **Editing the PDF digitally is prohibited except for signing the code of conduct and answering multiple choice questions**.
- Make sure that the **QR codes are visible** on every uploaded page. Otherwise, we cannot grade your submission.
- **You must solve the specified version of the problem**. Different problems may have different version: e.g. Problem 1 (Version A), Problem 5 (Version C), etc. If you solve the wrong version you get **zero** points.
- Only write on the provided sheets, **submitting your own additional sheets is not possible**.
- Last two pages can be used as scratch paper.
- All sheets (including scratch paper) have to be submitted to the upload queue. Missing pages will be considered empty.
- **Only use a black or blue color (no red or green)! Pencils are allowed.**
- Write your answers only in the provided solution boxes or the scratch paper.
- **For problems that say "Justify your answer" you only get points if you provide a valid explanation.**
- **For problems that say "derive" you only get points if you provide a valid derivation.**
- If a problem does not say "Justify your answer" or "derive", it's sufficient to only provide the correct answer.
- Instructor announcements and clarifications will be posted **on Piazza** with email notifications.
- Exercise duration - 75 minutes.

Problem 1 (Version C)

0 1 2 3 4 5 6

We compute the density using the change of variables formula

$$p_2(\boldsymbol{x}^{(0)}) = p_1(f^{-1}(\boldsymbol{x}^{(0)})) \left| \frac{\partial f^{-1}(\boldsymbol{x}^{(0)})}{\partial \boldsymbol{x}} \right| .$$

The inverse transformation is

$$f^{-1}(\boldsymbol{x}^{(0)}) = \boldsymbol{A}^{-1}\boldsymbol{x}^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/8 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \end{pmatrix} .$$

Therefore,

$$p_1(f^{-1}(\boldsymbol{x}^{(0)})) = \frac{1}{4}.$$

Since f is a linear transformation, the Jacobian determinant is

$$\left| \frac{\partial f^{-1}(\boldsymbol{x}^{(0)})}{\partial \boldsymbol{x}} \right| = \det(\boldsymbol{A}^{-1}) = 8.$$

Putting everything together, we get

$$p_2(\boldsymbol{x}^{(0)}) = \frac{1}{4} \cdot 8 = 2.$$

Problem 2 (Version A)

0
1
2
3
4
5
6

Our goal is to find a transformation $T_\phi : [0,1] \to \mathbb{R}$ such that

$$\begin{aligned} F_x(a) &= \Pr(x \leq a) \\ &= \Pr(T_\phi(u) \leq a) \\ &= \Pr(u \leq T_\phi^{-1}(a)) \\ &= F_u(T_\phi^{-1}(a)) \\ &= T_\phi^{-1}(a) \end{aligned}$$

where F_u is the CDF of the Uniform([0, 1]) distribution.
We can rewrite the above equation as

$$\begin{aligned} \frac{1}{1+\exp(-\phi a)} &= T_\phi^{-1}(a) \\ \frac{1}{1+\exp(-\phi T_\phi(a))} &= T_\phi^{-1}(T_\phi(a)) \\ \frac{1}{1+\exp(-\phi T_\phi(a))} &= a \\ \exp(-\phi T_\phi(a)) &= \frac{1}{a} - 1 \\ T_\phi(a) &= -\frac{1}{\phi} \log\left(\frac{1}{a} - 1\right) \end{aligned}$$

Hence, $T_\phi(a) = -\frac{1}{\phi} \log\left(\frac{1}{a} - 1\right)$ is the desired transformation.

Problem 3 (Version B)

0
1
2

a)

No, since $N \neq M$ means that the transformation f is not invertible. A normalizing flow model can only be defined using an invertible transformation.

0
1
2

b)

Yes, since there are no restrictions on the decoder in a VAE.We might need to apply some nonlinearity to ensure that the parameters are valid (e.g., nonnegative).

0
1
2

c)

Yes, since there are no restrictions on the the generator in a GAN.

Problem 4 (Version B)

0
1
2
3
4
5
6
7
8

We introduce a vector $q \in \{0,1\}^D$ of binary variables, indicating which input features are perturbed by the adversary.
We can then introduce $2 \cdot D$ constraints to express that $q_d = 0 \implies \tilde{x}_d = X_d$:

$$\tilde{x}_d - x_d \leq q_d \epsilon \; \forall i, j$$
$$\tilde{x}_d - x_d \geq q_d \epsilon \; \forall i, j$$

and one constraint to ensure that at most η pixels are perturbed:

$$\sum_{d=0}^{D-1} q_d \leq \eta$$

Problem 5 (Version B)

0 1 2

a)

The only root of its characteristic polynomial is 1 which is not strictly outside the unit circle. Therefore the process is not stationary.

0 1 2

b)

By plugging in the original process we see that

$$X'_t = (X_{t-1} + \varepsilon_t) - X_{t-1} = \varepsilon_t.$$

As such the sequence elements will just be i.i.d. noise variables which directly fulfill the definition of stationarity: their mean is 0 and therefore constant, their covariance is also 0 and thus independent of t and, finally, their variance is 1, so finite.

Problem 6 (Version D)

0
1
2
3
4
5
6
7
8

We are looking for the most likely latent state at a single point in time $\arg\max_{Z_2} \Pr(Z_2 \mid X_{1:3})$ which we can get from the forward-backward algorithm.

$$\Pr(Z_2 \mid X_{1:3}) \propto \Pr(Z_2, X_{1:3}) = \Pr(Z_2, X_{1:2}) \cdot \Pr(X_3 \mid Z_2) = \boldsymbol{\alpha}_2 \odot \boldsymbol{\beta}_2$$

We compute the forward and backward variables $\boldsymbol{\alpha}_t$ and $\boldsymbol{\beta}_t$ as follows.

$$\boldsymbol{\alpha}_1 = \boldsymbol{B}_{:,2} \odot \boldsymbol{\pi} \propto \begin{pmatrix} 2 \\ 6 \\ 0 \end{pmatrix} \qquad \boldsymbol{\alpha}_2 = \boldsymbol{B}_{:,2} \odot \boldsymbol{A}^{\mathsf{T}} \boldsymbol{\alpha}_1 \propto \boldsymbol{B}_{:,2} \odot \begin{pmatrix} 16 \\ 14 \\ 10 \end{pmatrix} \propto \begin{pmatrix} 16 \\ 42 \\ 0 \end{pmatrix}$$

$$\boldsymbol{\beta}_3 \propto \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad \boldsymbol{\beta}_2 = \boldsymbol{A}(\boldsymbol{B}_{:,3} \odot \boldsymbol{\beta}_3) \propto \boldsymbol{A} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \propto \begin{pmatrix} 9 \\ 7 \\ 13 \end{pmatrix}$$

In the end, we get that

$$\Pr(Z_2 \mid X_{1:3}) \propto \boldsymbol{\alpha}_2 \odot \boldsymbol{\beta}_2 = \begin{pmatrix} 144 \\ 294 \\ 0 \end{pmatrix}$$

and therefore the most likely latent state Z_2 is 2.

Problem 7 (Version D)

0 1 2 3 4 5 6

There are two equivalent ways to answer this question.

- We know that the inter-event times τ_i in a homogeneous Poisson process with rate are distributed according to the Exponential($\frac{1}{\mu}$) distribution. Therefore,

$$\Pr(t_1 > T) = \Pr(\tau_1 > T) = \exp(-\mu T).$$

- Equivalently, we know from the properties of the Poisson process that the number of events N follows Poisson($\int_0^T \mu dt$) = Poisson(μT). Hence, we can use the probability mass function of the Poisson distribution to compute

$$\Pr(N = 0) = \frac{(\mu T)^0 \exp(-\mu T)}{0!} = \exp(-\mu T).$$

Problem 8: Graphs - Clustering (Version A)

0 1 2 3 4 5 6

We compute the likelihood

$$P(A|\boldsymbol{\eta}, \mathbf{z}) = \prod_{ij} \frac{(\eta_{z_i z_j})^{A_{ij}}}{A_{ij}!} e^{-\eta_{z_i z_j}}$$

We can equivalently maximize the log-likelihood:

$$\begin{aligned}\log P(A|\boldsymbol{\eta}, \mathbf{z}) &= \sum_{ij} - \log(A_{ij}!) + A_{ij} \log \eta_{z_i z_j} - \eta_{z_i z_j} \\ &= -\sum_{ij} \log(A_{ij}!) + \sum_{ij} A_{ij} \log \eta_{z_i z_j} - \sum_{ij} \eta_{z_i z_j}\end{aligned}$$

We make use of the shorthand notation $N_p = \sum_{i=1}^{N} \mathbb{1}(z_i = p)$ and $M_{pq} = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} \mathbb{1}(z_i = p, z_j = q)$, and rewrite the log-likelihood:

$$\log P(A|\boldsymbol{\eta}, \mathbf{z}) = const + \sum_{p,q} m_{pq} \log \eta_{pq} - \sum_{p,q} n_p n_q \eta_{pq}.$$

We compute the derivative with respect to η_{pq} and set it to 0:

$$\frac{\partial \log P(A|\boldsymbol{\eta}, \mathbf{z})}{\partial \eta_{pq}} = \frac{m_{pq}}{\eta_{pq}} - n_p n_q$$

which gives $\eta_{pq} = \frac{m_{pq}}{n_p n_q}$.

Problem 8: Graphs - Clustering (Version B)

0 1 2 3 4 5 6

We compute the likelihood

$$P(A|\boldsymbol{\eta},\mathbf{z}) = \prod_{ij} \frac{(\eta_{z_i z_j})^{A_{ij}}}{A_{ij}!} e^{-\eta_{z_i z_j}}$$

We can equivalently maximize the log-likelihood:

$$\begin{aligned}\log P(A|\boldsymbol{\eta},\mathbf{z}) &= \sum_{ij} -\log(A_{ij}!) + A_{ij}\log \eta_{z_i z_j} - \eta_{z_i z_j} \\ &= -\sum_{ij}\log(A_{ij}!) + \sum_{ij} A_{ij}\log \eta_{z_i z_j} - \sum_{ij}\eta_{z_i z_j}\end{aligned}$$

We make use of the shorthand notation $N_p = \sum_{i=1}^{N} \mathbb{1}(z_i = p)$ and $M_{pq} = \sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}\mathbb{1}(z_i = p, z_j = q)$, and rewrite the log-likelihood:

$$\log P(A|\boldsymbol{\eta},\mathbf{z}) = const + \sum_{p,q} m_{pq}\log \eta_{pq} - \sum_{p,q} n_p n_q \eta_{pq}.$$

We compute the derivative with respect to η_{pq} and set it to 0:

$$\frac{\partial \log P(A|\boldsymbol{\eta},\mathbf{z})}{\partial \eta_{pq}} = \frac{m_{pq}}{\eta_{pq}} - n_p n_q$$

which gives $\eta_{pq} = \frac{m_{pq}}{n_p n_q}$.

Problem 8: Graphs - Clustering (Version C)

0
1
2
3
4
5
6

We compute the likelihood

$$P(A|\boldsymbol{\eta}, \boldsymbol{z}) = \prod_{ij} \frac{(\eta_{z_i z_j})^{A_{ij}}}{A_{ij}!} e^{-\eta_{z_i z_j}}$$

We can equivalently maximize the log-likelihood:

$$\begin{aligned}\log P(A|\boldsymbol{\eta}, \boldsymbol{z}) &= \sum_{ij} - \log(A_{ij}!) + A_{ij} \log \eta_{z_i z_j} - \eta_{z_i z_j} \\ &= - \sum_{ij} \log(A_{ij}!) + \sum_{ij} A_{ij} \log \eta_{z_i z_j} - \sum_{ij} \eta_{z_i z_j}\end{aligned}$$

We make use of the shorthand notation $N_p = \sum_{i=1}^{N} \mathbb{1}(z_i = p)$ and $M_{pq} = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} \mathbb{1}(z_i = p, z_j = q)$, and rewrite the log-likelihood:

$$\log P(A|\boldsymbol{\eta}, \boldsymbol{z}) = const + \sum_{p,q} m_{pq} \log \eta_{pq} - \sum_{p,q} n_p n_q \eta_{pq}.$$

We compute the derivative with respect to η_{pq} and set it to 0:

$$\frac{\partial \log P(A|\boldsymbol{\eta}, \boldsymbol{z})}{\partial \eta_{pq}} = \frac{m_{pq}}{\eta_{pq}} - n_p n_q$$

which gives $\eta_{pq} = \frac{m_{pq}}{n_p n_q}$.

Problem 8: Graphs - Clustering (Version D)

0
1
2
3
4
5
6

We compute the likelihood

$$P(A|\boldsymbol{\eta}, \boldsymbol{z}) = \prod_{ij} \frac{(\eta_{z_i z_j})^{A_{ij}}}{A_{ij}!} e^{-\eta_{z_i z_j}}$$

We can equivalently maximize the log-likelihood:

$$\begin{aligned} \log P(A|\boldsymbol{\eta}, \boldsymbol{z}) &= \sum_{ij} -\log(A_{ij}!) + A_{ij} \log \eta_{z_i z_j} - \eta_{z_i z_j} \\ &= -\sum_{ij} \log(A_{ij}!) + \sum_{ij} A_{ij} \log \eta_{z_i z_j} - \sum_{ij} \eta_{z_i z_j} \end{aligned}$$

We make use of the shorthand notation $N_p = \sum_{i=1}^{N} \mathbb{1}(z_i = p)$ and $M_{pq} = \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} \mathbb{1}(z_i = p, z_j = q)$, and rewrite the log-likelihood:

$$\log P(A|\boldsymbol{\eta}, \boldsymbol{z}) = const + \sum_{p,q} m_{pq} \log \eta_{pq} - \sum_{p,q} n_p n_q \eta_{pq}.$$

We compute the derivative with respect to η_{pq} and set it to 0:

$$\frac{\partial \log P(A|\boldsymbol{\eta}, \boldsymbol{z})}{\partial \eta_{pq}} = \frac{m_{pq}}{\eta_{pq}} - n_p n_q$$

which gives $\eta_{pq} = \frac{m_{pq}}{n_p n_q}$.

Problem 9: Graphs - Ranking (Version A)

a)

0
1
2
3

The stationnary distribution of the random walk associated with G is the vector $\pi(\infty) = [1, 0, 0, 0]$ satisfies $A\pi(\infty) = \pi(\infty)$ and normalized to 1.

b)

0
1

It is impossible to get from state 1 to other states i.e. state 1 is a dead end without out-links.

0
1
2
3
4
5
6

c)

The node 1 is a dead end. Therefore, there are three options to make the graph G' irreducible: add edge $(1, 2)$, $(1, 3)$ or $(1, 4)$.
The three systems of pagerank equations are respectively:

$$\begin{cases} r_1 = \dfrac{r_2}{2} + \dfrac{r_3}{2} \\ r_2 = r_4 + r_1 \\ r_3 = \dfrac{r_2}{2} \\ r_4 = \dfrac{r_3}{2} \end{cases} \qquad \begin{cases} r_1 = \dfrac{r_2}{2} + \dfrac{r_3}{2} \\ r_2 = r_4 \\ r_3 = \dfrac{r_2}{2} + r_1 \\ r_4 = \dfrac{r_3}{2} \end{cases} \qquad \begin{cases} r_1 = \dfrac{r_2}{2} + \dfrac{r_3}{2} \\ r_2 = r_4 \\ r_3 = \dfrac{r_2}{2} \\ r_4 = \dfrac{r_3}{2} + r_1 \end{cases}$$

where we also enforce $r_1 + r_2 + r_3 + r_4 = 1$. Solving the systems lead respectively to:

$$\begin{cases} r_1 = \dfrac{r_2}{2} + \dfrac{r_3}{2} \\ r_2 = \dfrac{4r_1}{3} \\ r_3 = \dfrac{2r_1}{3} \\ r_4 = \dfrac{r_1}{3} \end{cases} \qquad \begin{cases} r_1 = \dfrac{r_2}{2} + \dfrac{r_3}{2} \\ r_2 = \dfrac{2r_1}{3} \\ r_3 = \dfrac{4r_1}{3} \\ r_4 = \dfrac{2r_1}{3} \end{cases} \qquad \begin{cases} r_1 = \dfrac{r_2}{2} + \dfrac{r_3}{2} \\ r_2 = \dfrac{4r_1}{3} \\ r_3 = \dfrac{2r_1}{3} \\ r_4 = \dfrac{4r_1}{3} \end{cases}$$

Taking into account the normalization constraint, we obtain $r_1 = \frac{3}{10}, r_2 = \frac{4}{10}, r_3 = \frac{2}{10}, r_4 = \frac{1}{10}$ and $r_1 = \frac{3}{11}, r_2 = \frac{2}{11}, r_3 = \frac{4}{11}, r_4 = \frac{2}{11}$ and $r_1 = \frac{3}{13}, r_2 = \frac{4}{13}, r_3 = \frac{2}{13}, r_4 = \frac{4}{13}$.
The best edge to add is $(1, 2)$ to maximize the rank of node 1.

Problem 9: Graphs - Ranking (Version B)

a)

The stationnary distribution of the random walk associated with G is the vector $\pi(\infty) = [0, 1, 0, 0]$ satisfies $A\pi(\infty) = \pi(\infty)$ and normalized to 1. It defines the stationnary distribution of the random walk associated with G.

b)

It is impossible to get from state 2 to other states i.e. state 2 is a dead end without out-links.

c)

The node 2 is a dead end. Therefore, there are three options to make the graph G' irreducible: add edge $(2,1)$, $(2,3)$ or $(2,4)$.
The three systems of pagerank equations are respectively:

$$\begin{cases} r_1 = r_4 + r_2 \\ r_2 = \frac{r_1}{2} + \frac{r_3}{2} \\ r_3 = \frac{r_1}{2} \\ r_4 = \frac{r_3}{2} \end{cases} \quad \begin{cases} r_1 = r_4 \\ r_2 = \frac{r_1}{2} + \frac{r_3}{2} \\ r_3 = \frac{r_1}{2} + r_2 \\ r_4 = \frac{r_3}{2} \end{cases} \quad \begin{cases} r_1 = \frac{r_2}{2} + \frac{r_3}{2} \\ r_2 = \frac{r_1}{2} + \frac{r_3}{2} \\ r_3 = \frac{r_1}{2} \\ r_4 = \frac{r_3}{2} + r_2 \end{cases}$$

where we also enforce $r_1 + r_2 + r_3 + r_4 = 1$. Solving the systems lead respectively to:

$$\begin{cases} r_1 = \frac{4r_2}{3} \\ r_2 = \frac{r_1}{2} + \frac{r_3}{2} \\ r_3 = \frac{2r_2}{3} \\ r_4 = \frac{r_2}{3} \end{cases} \quad \begin{cases} r_1 = \frac{2r_2}{3} \\ r_2 = \frac{r_1}{2} + \frac{r_3}{2} \\ r_3 = \frac{4r_2}{3} \\ r_4 = \frac{2r_2}{3} \end{cases} \quad \begin{cases} r_1 = \frac{4r_2}{3} \\ r_2 = \frac{r_1}{2} + \frac{r_3}{2} \\ r_3 = \frac{2r_2}{3} \\ r_4 = \frac{4r_2}{3} \end{cases}$$

Taking into account the normalization constraint, we obtain $r_1 = \frac{4}{10}, r_2 = \frac{3}{10}, r_3 = \frac{2}{10}, r_4 = \frac{1}{10}$ and $r_1 = \frac{2}{11}, r_2 = \frac{3}{11}, r_3 = \frac{4}{11}, r_4 = \frac{2}{11}$ and $r_1 = \frac{4}{13}, r_2 = \frac{3}{13}, r_3 = \frac{2}{13}, r_4 = \frac{4}{13}$.
The best edge to add is $(2,1)$ to maximize the rank of node 1.

Problem 9: Graphs - Ranking (Version C)

a)

0
1
2
3

The stationnary distribution of the random walk associated with G is the vector $\pi(\infty) = [0, 0, 1, 0]$ satisfies $A\pi(\infty) = \pi(\infty)$ and normalized to 1. It defines the stationnary distribution of the random walk associated with G.

b)

0
1

It is impossible to get from state 3 to other states i.e. state 3 is a dead end without out-links.

0
1
2
3
4
5
6

c)

The node 3 is a dead end. Therefore, there are three options to make the graph G' irreducible: add edge $(3,1)$, $(3,2)$ or $(3,4)$.
The three systems of pagerank equations are respectively:

$$\begin{cases} r_1 = \dfrac{r_2}{2} + r_3 \\ r_2 = r_4 \\ r_3 = \dfrac{r_1}{2} + \dfrac{r_2}{2} \\ r_4 = \dfrac{r_1}{2} \end{cases} \qquad \begin{cases} r_1 = \dfrac{r_2}{2} \\ r_2 = r_4 + r_3 \\ r_3 = \dfrac{r_1}{2} + \dfrac{r_2}{2} \\ r_4 = \dfrac{r_1}{2} \end{cases} \qquad \begin{cases} r_1 = \dfrac{r_2}{2} \\ r_2 = r_4 \\ r_3 = \dfrac{r_1}{2} + \dfrac{r_2}{2} \\ r_4 = \dfrac{r_1}{2} + r_3 \end{cases}$$

where we also enforce $r_1 + r_2 + r_3 + r_4 = 1$. Solving the systems lead respectively to:

$$\begin{cases} r_1 = \dfrac{4r_3}{3} \\ r_2 = \dfrac{2r_3}{3} \\ r_3 = \dfrac{r_1}{2} + \dfrac{r_2}{2} \\ r_4 = \dfrac{2r_3}{3} \end{cases} \qquad \begin{cases} r_1 = \dfrac{2r_3}{3} \\ r_2 = \dfrac{4r_3}{3} \\ r_3 = \dfrac{r_1}{2} + \dfrac{r_3}{2} \\ r_4 = \dfrac{1r_3}{3} \end{cases} \qquad \begin{cases} r_1 = \dfrac{2r_3}{3} \\ r_2 = \dfrac{4r_3}{2} \\ r_3 = \dfrac{r_1}{2} + \dfrac{r_3}{2} \\ r_4 = \dfrac{4r_3}{3} \end{cases}$$

Taking into account the normalization constraint, we obtain $r_1 = \frac{4}{11}, r_2 = \frac{2}{11}, r_3 = \frac{3}{11}, r_4 = \frac{2}{11}$ and $r_1 = \frac{2}{10}, r_2 = \frac{4}{10}, r_3 = \frac{3}{10}, r_4 = \frac{1}{10}$ and $r_1 = \frac{2}{13}, r_2 = \frac{4}{13}, r_3 = \frac{3}{13}, r_4 = \frac{4}{13}$.
The best edge to add is $(3,1)$ to maximize the rank of node 1.

Problem 9: Graphs - Ranking (Version D)

a)

0
1
2
3

The stationnary distribution of the random walk associated with G is the vector $\pi(\infty) = [0, 0, 0, 1]$ satisfies $A\pi(\infty) = \pi(\infty)$ and normalized to 1. It defines the stationnary distribution of the random walk associated with G.

b)

0
1

It is impossible to get from state 4 to other states i.e. state 4 is a dead end without out-links.

0
1
2
3
4
5
6

c)

The node 4 is a dead end. Therefore, there are three options to make the graph G' irreducible: add edge $(4,1)$, $(4,2)$ or $(4,3)$.
The three systems of pagerank equations are respectively:

$$\begin{cases} r_1 = \frac{r_3}{2} + r_4 \\ r_2 = r_1 \\ r_3 = \frac{r_2}{2} \\ r_4 = \frac{r_2}{2} + \frac{r_3}{2} \end{cases} \quad \begin{cases} r_1 = \frac{r_3}{2} \\ r_2 = r_1 + r_4 \\ r_3 = \frac{r_2}{2} \\ r_4 = \frac{r_2}{2} + \frac{r_3}{2} \end{cases} \quad \begin{cases} r_1 = \frac{r_3}{2} \\ r_2 = r_1 \\ r_3 = \frac{r_2}{2} + r_4 \\ r_4 = \frac{r_2}{2} + \frac{r_3}{2} \end{cases}$$

where we also enforce $r_1 + r_2 + r_3 + r_4 = 1$. Solving the systems lead respectively to:

$$\begin{cases} r_1 = \frac{4r_4}{3} \\ r_2 = \frac{4r_4}{3} \\ r_3 = \frac{2r_4}{3} \\ r_4 = \frac{r_2}{2} + \frac{r_3}{2} \end{cases} \quad \begin{cases} r_1 = \frac{1r_4}{3} \\ r_2 = \frac{4r_4}{3} \\ r_3 = \frac{2r_4}{3} \\ r_4 = \frac{r_2}{2} + \frac{r_3}{2} \end{cases} \quad \begin{cases} r_1 = \frac{2r_4}{3} \\ r_2 = \frac{2r_4}{3} \\ r_3 = \frac{4r_4}{3} \\ r_4 = \frac{r_2}{2} + \frac{r_3}{2} \end{cases}$$

Taking into account the normalization constraint, we obtain $r_1 = \frac{4}{13}, r_2 = \frac{4}{13}, r_3 = \frac{2}{13}, r_4 = \frac{3}{13}$ and $r_1 = \frac{1}{10}, r_2 = \frac{4}{10}, r_3 = \frac{2}{10}, r_4 = \frac{3}{10}$ and $r_1 = \frac{2}{11}, r_2 = \frac{2}{11}, r_3 = \frac{4}{11}, r_4 = \frac{3}{11}$.
The best edge to add is $(3,1)$ to maximize the rank of node 1.

Problem 10: Graphs - Semi-Supervised Learning (Version A)

a)

This problem is equivalent to the minimum cut problem separating the two clusters.

Optimal label assignments are $\boldsymbol{y}_U = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\boldsymbol{y}_U = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. They achieve a minimum cost of 2.

b)

We first write the Laplacian L and L' in block form of both graphs G and G':

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{L}_{SS} & \boldsymbol{L}_{SU} \\ \boldsymbol{L}_{US} & \boldsymbol{L}_{UU} \end{bmatrix} \in \mathbb{R}^{5\times 5}$$

and

$$\boldsymbol{L}' = \begin{bmatrix} \boldsymbol{L}'_{SS} & \boldsymbol{L}'_{SU} \\ \boldsymbol{L}'_{US} & \boldsymbol{L}'_{UU} \end{bmatrix} \in \mathbb{R}^{6\times 6}$$

Given this notation, we can write the closed-form solution for both graphs:

$$\boldsymbol{y}^*_U = -\boldsymbol{L}_{UU}^{-1}\boldsymbol{L}_{US}\hat{\boldsymbol{y}}_S \text{ and } \boldsymbol{y}'^*_U = -\boldsymbol{L}'^{-1}_{UU}\boldsymbol{L}'_{US}\hat{\boldsymbol{y}}'_S$$

We note that $\boldsymbol{L}'_{UU} = \boldsymbol{L}_{UU}$, $\boldsymbol{L}'_{SU} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{L}_{SU} \end{bmatrix}$ and $\hat{\boldsymbol{y}}'_S = \begin{bmatrix} 1 \\ \hat{\boldsymbol{y}}_S \end{bmatrix}$.

Finally, we obtain:

$$\begin{aligned} \boldsymbol{y}'^*_U &= -\boldsymbol{L}'^{-1}_{UU}\boldsymbol{L}'_{US}\hat{\boldsymbol{y}}'_S \\ &= -\boldsymbol{L}'^{-1}_{UU}(\boldsymbol{0} \times 1 + \boldsymbol{L}_{SU}\hat{\boldsymbol{y}}_S) \\ &= \boldsymbol{y}^*_U \end{aligned}$$

and the final solution is $\boldsymbol{y}'^* = \begin{bmatrix} 1 \\ \boldsymbol{y}^* \end{bmatrix}$.

Problem 10: Graphs - Semi-Supervised Learning (Version B)

0 1 2 3

a)

This problem is equivalent to the minimum cut problem separating the two clusters.

Optimal label assignments are $\boldsymbol{y}_U = \begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\boldsymbol{y}_U = \begin{bmatrix}0\\0\\0\end{bmatrix}$. They achieve a minimum cost of 2.

0 1 2 3 4 5

b)

We first write the Laplacian L and L' in block form of both graphs G and G':

$$\boldsymbol{L} = \begin{bmatrix}\boldsymbol{L}_{SS} & \boldsymbol{L}_{SU}\\ \boldsymbol{L}_{US} & \boldsymbol{L}_{UU}\end{bmatrix} \in \mathbb{R}^{5\times 5}$$

and

$$\boldsymbol{L}' = \begin{bmatrix}\boldsymbol{L}'_{SS} & \boldsymbol{L}'_{SU}\\ \boldsymbol{L}'_{US} & \boldsymbol{L}'_{UU}\end{bmatrix} \in \mathbb{R}^{6\times 6}$$

Given this notation, we can write the closed-form solution for both graphs:

$$\boldsymbol{y}^*_U = -\boldsymbol{L}^{-1}_{UU}\boldsymbol{L}_{US}\hat{\boldsymbol{y}}_S \text{ and } \boldsymbol{y}'^*_U = -\boldsymbol{L}'^{-1}_{UU}\boldsymbol{L}'_{US}\hat{\boldsymbol{y}}'_S$$

We note that $\boldsymbol{L}'_{UU} = \boldsymbol{L}_{UU}$, $\boldsymbol{L}'_{SU} = \begin{bmatrix}\boldsymbol{0} & \boldsymbol{L}_{SU}\end{bmatrix}$ and $\hat{\boldsymbol{y}}'_S = \begin{bmatrix}1\\ \hat{\boldsymbol{y}}_S\end{bmatrix}$.

Finally, we obtain:

$$\begin{aligned}\boldsymbol{y}'^*_U &= -\boldsymbol{L}'^{-1}_{UU}\boldsymbol{L}'_{US}\hat{\boldsymbol{y}}'_S\\ &= -\boldsymbol{L}'^{-1}_{UU}(\boldsymbol{0}\times 1 + \boldsymbol{L}_{SU}\hat{\boldsymbol{y}}_S)\\ &= \boldsymbol{y}^*_U\end{aligned}$$

and the final solution is $\boldsymbol{y}'^* = \begin{bmatrix}1\\ \boldsymbol{y}^*\end{bmatrix}$.

Problem 10: Graphs - Semi-Supervised Learning (Version C)

a)

0 1 2 3

This problem is equivalent to the minimum cut problem separating the two clusters.

Optimal label assignments are $\boldsymbol{y}_U = \begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\boldsymbol{y}_U = \begin{bmatrix}0\\0\\0\end{bmatrix}$. They achieve a minimum cost of 2.

b)

0 1 2 3 4 5

We first write the Laplacian L and L' in block form of both graphs G and G':

$$\boldsymbol{L} = \begin{bmatrix}\boldsymbol{L}_{SS} & \boldsymbol{L}_{SU}\\ \boldsymbol{L}_{US} & \boldsymbol{L}_{UU}\end{bmatrix} \in \mathbb{R}^{5\times 5}$$

and

$$\boldsymbol{L}' = \begin{bmatrix}\boldsymbol{L}'_{SS} & \boldsymbol{L}'_{SU}\\ \boldsymbol{L}'_{US} & \boldsymbol{L}'_{UU}\end{bmatrix} \in \mathbb{R}^{6\times 6}$$

Given this notation, we can write the closed-form solution for both graphs:

$$\boldsymbol{y}^*_U = -\boldsymbol{L}^{-1}_{UU}\boldsymbol{L}_{US}\hat{\boldsymbol{y}}_S \text{ and } \boldsymbol{y}'^*_U = -\boldsymbol{L}'^{-1}_{UU}\boldsymbol{L}'_{US}\hat{\boldsymbol{y}}'_S$$

We note that $\boldsymbol{L}'_{UU} = \boldsymbol{L}_{UU}$, $\boldsymbol{L}'_{SU} = \begin{bmatrix}\boldsymbol{0} & \boldsymbol{L}_{SU}\end{bmatrix}$ and $\hat{\boldsymbol{y}}'_S = \begin{bmatrix}1\\ \hat{\boldsymbol{y}}_S\end{bmatrix}$.
Finally, we obtain:

$$\begin{aligned}\boldsymbol{y}'^*_U &= -\boldsymbol{L}'^{-1}_{UU}\boldsymbol{L}'_{US}\hat{\boldsymbol{y}}'_S\\ &= -\boldsymbol{L}'^{-1}_{UU}(\boldsymbol{0}\times 1 + \boldsymbol{L}_{SU}\hat{\boldsymbol{y}}_S)\\ &= \boldsymbol{y}^*_U\end{aligned}$$

and the final solution is $\boldsymbol{y}'^* = \begin{bmatrix}1\\ \boldsymbol{y}^*\end{bmatrix}$.

Problem 10: Graphs - Semi-Supervised Learning (Version D)

0 1 2 3

a)

This problem is equivalent to the minimum cut problem separating the two clusters.

Optimal label assignments are $\boldsymbol{y}_U = \begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\boldsymbol{y}_U = \begin{bmatrix}0\\0\\0\end{bmatrix}$. They achieve a minimum cost of 2.

0 1 2 3 4 5

b)

We first write the Laplacian L and L' in block form of both graphs G and G':

$$\boldsymbol{L} = \begin{bmatrix}\boldsymbol{L}_{SS} & \boldsymbol{L}_{SU}\\ \boldsymbol{L}_{US} & \boldsymbol{L}_{UU}\end{bmatrix} \in \mathbb{R}^{5\times 5}$$

and

$$\boldsymbol{L}' = \begin{bmatrix}\boldsymbol{L}'_{SS} & \boldsymbol{L}'_{SU}\\ \boldsymbol{L}'_{US} & \boldsymbol{L}'_{UU}\end{bmatrix} \in \mathbb{R}^{6\times 6}$$

Given this notation, we can write the closed-form solution for both graphs:

$$\boldsymbol{y}^*_U = -\boldsymbol{L}^{-1}_{UU}\boldsymbol{L}_{US}\hat{\boldsymbol{y}}_S \text{ and } \boldsymbol{y}'^*_U = -\boldsymbol{L}'^{-1}_{UU}\boldsymbol{L}'_{US}\hat{\boldsymbol{y}}'_S$$

We note that $\boldsymbol{L}'_{UU} = \boldsymbol{L}_{UU}$, $\boldsymbol{L}'_{SU} = \begin{bmatrix}\boldsymbol{0} & \boldsymbol{L}_{SU}\end{bmatrix}$ and $\hat{\boldsymbol{y}}'_S = \begin{bmatrix}1\\ \hat{\boldsymbol{y}}_S\end{bmatrix}$.

Finally, we obtain:

$$\begin{aligned}\boldsymbol{y}'^*_U &= -\boldsymbol{L}'^{-1}_{UU}\boldsymbol{L}'_{US}\hat{\boldsymbol{y}}'_S\\ &= -\boldsymbol{L}'^{-1}_{UU}(\boldsymbol{0} \times 1 + \boldsymbol{L}_{SU}\hat{\boldsymbol{y}}_S)\\ &= \boldsymbol{y}^*_U\end{aligned}$$

and the final solution is $\boldsymbol{y}'^* = \begin{bmatrix}1\\ \boldsymbol{y}^*\end{bmatrix}$.

Additional space for solutions–clearly mark the (sub)problem your answers are related to and strike out invalid solutions.