

Tutorials: NDMI025 - Randomized Algorithms

Karel Král

March 18, 2021

This text is a work in progress, do not distribute. All errors in this text are on purpose. Please report them to my email kralka@iuuk.mff.cuni...

Contents

1 Exercises	5
1.1 Tutorial 1.....	5
1.2 Tutorial 2.....	7
2 Theory	9
2.1 Probability 101.....	9
2.2 Markov Chain.....	9
3 Solutions	11
3.1 Tutorial 1.....	11
3.2 Tutorial 2.....	20

Chapter 1

Exercises

1.1 Tutorial

1.
 - Can you all hear me?
 - If you are uncomfortable asking a question in English, just ask in Czech/Slovak and I will translate.
 - Have you all taken:
 - (a) a probability course (discrete probability, random variables, expected value, variance, Markov, Chernoff)
 - (b) a linear algebra course (matrix operations, linear maps, eigenvectors and eigenvalues, discriminant)
 - (c) a graph theory course (what a combinatorial graph is, bipartite, complete, coloring)
 - (d) a combinatorics course (factorial, binomial coefficients)
 - (e) an algorithms / programming course (big-O notation, possibly understanding Python based on the other question)
 - This class is heavy on theory. Are you interested in computer simulations and or implementations? If so:
 - (a) Python
 - (b) R
 - (c) C++

Solution: 1

2. You are presented with two sealed envelopes. There are k € in one of those and ℓ € in the other ($k, \ell \in \mathbb{N}$ but you do not know k, ℓ in advance). You may open an envelope and (based on what you see) decide to take this one or the other (without looking into both).
 - (a) Is there a way how to walk away with the larger amount of money with probability strictly larger than 0.5?
 - (b) What is the expected value you walk away with (in terms of k, ℓ)?
 - (c) Simulate.

Solution: 2

3. Graph isomorphism. You have seen an interactive proof of graph non-isomorphism on the class. Can you come up with an interactive proof of graph isomorphism?

Solution: [3](#)

4. We will focus on random walks and their properties a lot.
- Random walks are useful when analysing algorithms – “two coloring without monochromatic triangle” of three-colorable graph.
 - Random numbers in the computer are often expensive to generate, can we reduce number of used random bits (expanders)? Or even get a deterministic algorithm?
 - To sample from extremely large spaces.

Let $n \in \mathbb{N}$, say $n = 30$. Let us the following problem we start with $X_0 = \lfloor n/2 \rfloor$ and do the following process:

- if $X_i \in \{0, n\}$ we stop
- we set $X_{i+1} = X_i + \delta$ where δ is picked uniformly at random from $\{-1, 1\}$

- Is this a Markov chain (Definition [2.2](#))? If so can you write it’s matrix?
- What is the expected number of steps until stopping?

Solution: [4](#)

5. Think of some example MCs.
- Create a MC that is irreducible.
 - Create a MC that is not irreducible.
 - Create a MC that is periodic.
 - Create a MC that is not periodic.
 - Compute a stationary distribution of the following MC:

$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

- Create a MC that has more stationary distributions.

Solution: [5](#)

6. We are collectors and we want to collect all n kinds of coupons. Coupons are sold in packages which all look the same. Thus when we buy an coupon, we buy one of n kinds uniformly at random. This is known as the *coupon collector* problem.

- What is the expected number of coupons we need to buy to get all kinds?
- How many coupons do we need to buy to have probability at least $1 - q$ of collecting all kinds?
- What is the Markov chain? Is this similar to a random walk on some graph?
- Simulate.

Solution: [6](#)

1.2 Tutorial

1. Find a family of oriented graphs of constant in-degree and constant out-degree and as large hitting time as possible.

Note that similar situation could happen on undirected graphs where the probabilities of traversing edge one way and the other way would not be the same. Which is in principle almost an oriented graph.

Solution: [1](#)

2. Let $A \in \mathbb{R}^{n \times n}$ be a matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. Show that the matrix $A + dI_n$ has eigenvalues $d + \lambda_1, \dots, d + \lambda_n$.

Solution: [2](#)

3. Show Courant-Fisher: Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix ($A^T = A$). Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be its eigenvalues. Show

(a) $\lambda_1 = \max_{x \in \mathbb{R}^n, \|x\|=1} x^T A x$

(b) $\lambda_n = \min_{x \in \mathbb{R}^n, \|x\|=1} x^T A x$

- (c) The eigenvalue λ_2 can be computed similarly $\lambda_2 = \max_{x \in \mathbb{R}^n, \|x\|=1, x^T u_1 = 0} x^T A x$ (where u_1 is the eigenvector corresponding to λ_1). We can get other eigenvalues in a similar manner. Moreover we could use this to prove the interlacing theorem. See https://en.wikipedia.org/wiki/Min-max_theorem

Solution: [3](#)

4. Show that a connected d -regular graph is bipartite iff the least eigenvalue of its adjacency matrix is $-d$.

Solution: [4](#)

5. Compute the eigenvalues and eigenvectors of the following graphs:

(a) K_n , the complete graph on n vertices.

(b) $K_{n,n}$, the complete bipartite graph with partites of size n each.

(c) C_n , the cycle on n vertices.

Solution: [5](#)

6. Distributed discrete logarithm algorithm.

Solution: [6](#)

Chapter 2

Theory

2.1 Probability 101

Probability 101

2.2 Markov Chain

Definition. A discrete-time Markov chain is a sequence of random variables X_0, X_1, X_2, \dots with the Markov property:

$$\Pr[X_{n+1} = x \mid X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] = \Pr[X_{n+1} = x \mid X_n = x_n]$$

(if both are defined, i.e., $\Pr[X_0 = x_0, X_1 = x_1, \dots, X_n = x_n] > 0$)

and the possible values of X_i form a countable set called the state space of the Markov chain.

The Markov property states that the process has no memory – the next state depends only on the current state. We will deal with a special case where the state space of each random variable will be the same and finite. Moreover we will deal with time-homogenous Markov chains, that is $\Pr[X_{n+1} \mid X_n] = \Pr[X_n \mid X_{n-1}]$ (the transition probabilities are time independent). Thus we will represent Markov chains by their transition matrices – if a Markov chain has n states its transition matrix is $P \in [0, 1]^{n \times n}$ such that $P_{i,j} = \Pr[X_{n+1} = i \mid X_n = j]$ (thus column sums are equal to one).

If we take a probability distribution $\pi \in [0, 1]^n$ and multiply it by the transition matrix we get the probability distribution after one step $P\pi$.

There are several interesting properties of Markov chains:

- We say that a MC is *irreducible* iff for each pair of states $i, j \in [n]$ there is a time $k \in \mathbb{N}$ such that $(P^k)_{i,j} > 0$ (we can get from any state to any state).
- We say that a MC is *periodic* iff there is a state $i \in [n]$ and a period $p \in \mathbb{N}, p > 1$ such that for each time $k \in \mathbb{N}$ we have $(P^k)_{i,i} > 0 \Rightarrow p \mid k$ that is probability of staying at state i is positive only for multiples of the period.
- We say that $\pi \in [0, 1]^n$ is a *stationary distribution* of a given MC iff $P\pi = \pi$ (the distribution is the same after one step).

Theorem 1. If a MC is aperiodic and irreducible it has a unique stationary distribution π . Moreover for all pairs of states $i, j \in [n]$ we know that

$$\lim_{t \rightarrow \infty} (P^t)_{i,j} = \pi_i$$

Chapter 3

Solutions

3.1 Tutorial

1.

- Can you all hear me?
- If you are uncomfortable asking a question in English, just ask in Czech/Slovak and I will translate.
- Have you all taken:
 - (a) a probability course (discrete probability, random variables, expected value, variance, Markov, Chernoff)
 - (b) a linear algebra course (matrix operations, linear maps, eigenvectors and eigenvalues, discriminant)
 - (c) a graph theory course (what a combinatorial graph is, bipartite, complete, coloring)
 - (d) a combinatorics course (factorial, binomial coefficients)
 - (e) an algorithms / programming course (big-O notation, possibly understanding Python based on the other question)
- This class is heavy on theory. Are you interested in computer simulations and or implementations? If so:
 - (a) Python
 - (b) R
 - (c) C++

2. You are presented with two sealed envelopes. There are k € in one of those and ℓ € in the other ($k, \ell \in \mathbb{N}$ but you do not know k, ℓ in advance). You may open an envelope and (based on what you see) decide to take this one or the other (without looking into both).

- (a) Is there a way how to walk away with the larger amount of money with probability strictly larger than 0.5?

Solution: Pick an envelope uniformly at random. If you see m € toss a fair coin until you get Tails. If the number of tosses was strictly less than m keep the envelope, otherwise take the other. If $k < \ell$ then the probability of keeping the envelope with k € is strictly less than the probability of keeping the envelope with ℓ €.

- (b) What is the expected value you walk away with (in terms of k, ℓ)?

Solution: Let us recall the sum of geometric series:

$$\begin{aligned} S &= \sum_{j=0}^n q^j \\ &= 1 + q + q^2 + \dots + q^n \\ &= 1 + q(1 + q + q^2 + \dots + q^{n-1}) \\ &= 1 + q(S - q^n) \end{aligned}$$

thus

$$\begin{aligned} S &= 1 + q(S - q^n) \\ S - qS &= 1 - q^{n+1} \\ S &= \frac{1 - q^{n+1}}{1 - q} \end{aligned} \quad (\text{pokud } q \neq 1)$$

and for the infinite case:

$$\begin{aligned} \sum_{j=0}^{\infty} q^j &= \lim_{n \rightarrow \infty} \sum_{j=0}^n q^j \\ &= \lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} \\ &= \frac{1}{1 - q} \end{aligned} \quad (\text{pokud } |q| < 1)$$

Thus exactly n tosses have probability for the general case where Tails has probability p and Heads has probability $1 - p$:

$$\Pr[n \text{ tosses}] = (1 - p)^{n-1}p \quad (\text{for any } n \in \mathbb{N}^+)$$

Probability of at most n tosses:

$$\begin{aligned} \Pr[1, 2, \dots, n \text{ tosses}] &= \sum_{j=1}^n p(1 - p)^{j-1} \\ &= p \sum_{j=1}^n (1 - p)^{j-1} \\ &= p \frac{1 - (1 - p)^n}{1 - (1 - p)} \end{aligned}$$

$$= 1 - (1 - p)^n$$

Probability that we keep k € (fair coin):

$$\begin{aligned} \Pr[\text{tosses} < k] &= \sum_{j=1}^{k-1} 0.5^j \\ &= 1 - 0.5^{k-1} \end{aligned}$$

Thus probability of walking away with k € is

$$\begin{aligned} \Pr[\text{winning } k\text{€}] &= \frac{1}{2}(1 - 0.5^{k-1}) + \frac{1}{2}0.5^{\ell-1} \\ &= \frac{1}{2} - 0.5^k + 0.5^\ell \\ &= \frac{1}{2} + (0.5^\ell - 0.5^k) \end{aligned}$$

Thus the expected win is

$$\mathbb{E}[\text{win}] = k \left(\frac{1}{2} + (0.5^\ell - 0.5^k) \right) + \ell \left(\frac{1}{2} + (0.5^k - 0.5^\ell) \right)$$

(c) Simulate.

Solution:

```
# https://docs.python.org/3/library/random.html
# Do not use for cryptography!
from random import randint
from random import random

def geometric(pr: float = 0.5) -> int:
    """pr is success probability, return the number of tosses until
    the first success."""
    assert pr > 0
    sample = 1
    fail_pr = 1 - pr
    while random() < fail_pr:
        sample += 1
    return sample

# Our unknown amounts.
envelopes = [5, 10]

N = 1000000          # Number of samples.
total_amount = 0    # Total sum that we got during all samples.
got_larger = 0      # Number of times we walked away with the larger sum.

for _ in range(N):
    # Pick the first envelope at random.
    chosen = randint(0, 1)
```

```
if geometric() < envelopes[chosen]:
    # Keep this one.
    pass
else:
    # Choose the other.
    chosen = 1 - chosen
if envelopes[chosen] >= envelopes[1 - chosen]:
    got_larger += 1
total_amount += envelopes[chosen]

k = envelopes[0]
l = envelopes[1]
pr_larger = 0.5 + abs(0.5**k - 0.5**l)
e_win = k * (0.5 + (0.5**l - 0.5**k)) + l * (0.5 + (0.5**k - 0.5**l))

print(f'Pr[selected larger] = {got_larger / N} (={pr_larger})')
print(f'E[win] = {total_amount / N} (={e_win})')

# Possible outcome:
# Pr[selected larger] = 0.529865 (=0.5302734375)
# E[win] = 7.649325 (=7.6513671875)
```

3. **Graph isomorphism.** You have seen an interactive proof of graph non-isomorphism on the class. Can you come up with an interactive proof of graph isomorphism?

Solution:

- Both the prover P and the verifier V know two graphs G_1, G_2 .
- The prover knows an isomorphism π such that $\pi(G_1) = G_2$. Formally $\pi: V(G_1) \rightarrow V(G_2)$ such that

$$(u, v) \in E(G_1) \Leftrightarrow (\pi(u), \pi(v)) \in E(G_2).$$

And by $\pi(G_1)$ we mean the graph $(\pi(V(G_1)), \{(\pi(u), \pi(v)) \mid (u, v) \in E(G_1)\})$.

- For ease of presentation we set $V(G_1) = V(G_2) = [n] = \{1, 2, 3, \dots, n\}$.
- The prover picks uniformly random permutation $\sigma \in S_n$ and sends the graph $G = \sigma(G_1)$.
- The verifier picks uniformly random number $i \in \{1, 2\}$ and asks verifier to present a permutation τ such that $\tau(G) = G_i$.
- If $i = 1$ then the prover sends $\tau = \sigma^{-1}$. If $i = 2$ then the prover sends $\tau = (\sigma \circ \pi)^{-1}$.

This is indeed an interactive proof:

- If the prover knows the isomorphism π , then all answers are correct.
- If G_1, G_2 are not isomorphic, then the verifier will pick a graph (either G_1 or G_2) which is not isomorphic with G with probability $1/2$.

Again the prover learns nothing about the isomorphism. If you find these interactive proofs interesting, take a look at “Zero Knowledge Proofs”.

Also note that our prover can be implemented efficiently as opposed to the case of graph non-isomorphism. In fact in some sense the prover proves that it knows the isomorphism (this can be made formal, see “Zero Knowledge Proofs of Knowledge”).

It is natural to repeat this protocol more times in order to boost the probabilities. This is called probability amplification. We will investigate this much more during the semester.

4. We will focus on random walks and their properties a lot.
- (a) Random walks are useful when analysing algorithms – “two coloring without monochromatic triangle” of three-colorable graph.
 - (b) Random numbers in the computer are often expensive to generate, can we reduce number of used random bits (expanders)? Or even get a deterministic algorithm?
 - (c) To sample from extremely large spaces.

Let $n \in \mathbb{N}$, say $n = 30$. Let us the following problem we start with $X_0 = \lfloor n/2 \rfloor$ and do the following process:

- if $X_i \in \{0, n\}$ we stop
 - we set $X_{i+1} = X_i + \delta$ where δ is picked uniformly at random from $\{-1, 1\}$
- (a) Is this a Markov chain (Definition [2.2](#))? If so can you write it’s matrix?

Solution: Yes (see the lecture video).

- (b) What is the expected number of steps until stopping?

Solution: Let us set

$$S_k = \mathbb{E}[\text{number of steps untill stopping, when starting at } k]$$

We know the following:

$$\begin{aligned} S_0 = S_n &= 0 \\ S_k &= 1 + \frac{1}{2}(S_{k-1} + S_{k+1}) \end{aligned} \quad (\text{by linearity of expectation})$$

The above is so-called difference equation. It is not terribly complicated, but not super easy to solve (hint try to consider equations for $d(k) = S_k - S_{k-1}$ to get rid of the “1+” term). You may look at https://en.wikipedia.org/wiki/Recurrence_relation Luckily when dealing with asymptotics thus we do not need exact estimates. And you will see some nice theoretical results tomorrow.

But it can be shown that

$$S_k = k(n - k)$$

which we can easily check that this is indeed a solution (note that we would also need that this is a unique solution, see solution methods on Wikipedia for this part):

$$\begin{aligned} S_k &= 1 + \frac{1}{2}(S_{k-1} + S_{k+1}) \\ S_k &= 1 + \frac{1}{2}((k-1)(n-(k-1)) + (k+1)(n-(k+1))) \\ S_k &= 1 + \frac{1}{2}((k-1)n - (k-1)^2 + (k+1)n - (k+1)^2) \\ S_k &= 1 + \frac{1}{2}(2kn - (k-1)^2 - (k+1)^2) \\ S_k &= 1 + \frac{1}{2}(2kn - 2k^2 - 2) \\ S_k &= k(n - k) \end{aligned}$$

5. Think of some example MCs.

(a) Create a MC that is irreducible.

Solution: Two states:

$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

(with probability 1/2 stay at the current state, with probability 1/2 switch to the other state).

(b) Create a MC that is not irreducible.

Solution: Two states:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

(always stay at the first state or immediately go there).

(c) Create a MC that is periodic.

Solution: Three states:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

(from the first state go always to the third, from the second always to the first and from the third always to the second).

(d) Create a MC that is not periodic.

Solution: Two states:

$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

(with probability 1/2 stay at the current state, with probability 1/2 switch to the other state).

(e) Compute a stationary distribution of the following MC:

$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

Solution: One eigenvalue is 1, the only stationary distribution $(1/2, 1/2)^T$. The other eigenvalue is 0 with the corresponding eigenvector $(1, -1)^T$ (this is not a distribution).

(f) Create a MC that has more stationary distributions.

Solution: Two states:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(always stay where we are).

6. We are collectors and we want to collect all n kinds of coupons. Coupons are sold in packages which all look the same. Thus when we buy an coupon, we buy one of n kinds uniformly at random. This is known as the *coupon collector* problem.

- (a) What is the expected number of coupons we need to buy to get all kinds?

Solution: Let t_i be the time to collect the i -th coupon kind after we have collected $i - 1$ coupons. The probability of buying the i -th coupon is

$$\Pr[\text{getting } i\text{-th coupon when already having } i - 1 \text{ coupons}] = \frac{n - (i - 1)}{n}$$

Thus t_i has geometric distribution (we are tossing the same probability and waiting for the first success). The expected value of t_i is:

$$\mathbb{E}[t_i] = \frac{n}{n - (i - 1)}$$

By linearity of expectation:

$$\begin{aligned} \mathbb{E}[\text{collecting}] &= \mathbb{E}[t_1 + t_2 + \dots + t_n] \\ &= \mathbb{E}[t_1] + \mathbb{E}[t_2] + \dots + \mathbb{E}[t_n] \\ &= \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \dots + \frac{n}{n-(n-1)} \\ &= nH_n \\ &= n \log(n) + n \cdot 0.577\dots + 1/2 + \mathcal{O}(1/n) \quad (\text{source Wikipedia}) \end{aligned}$$

- (b) How many coupons do we need to buy to have probability at least $1 - q$ of collecting all kinds?

Solution: We can use Markov inequality $\Pr[T > nH_n/q] \leq q$ (here T is the random variable telling us how many tosses are necessary).

- (c) What is the Markov chain? Is this similar to a random walk on some graph?

Solution: There might be more Markov chains corresponding to this problem. The states could be all subsets of $[n] = \{1, 2, 3, \dots, n\}$ (too big – not that nice to work with) or how many coupons have we collected so far (much smaller).

This corresponds to the cover time of a complete graph (when we have loops in each vertex).

- (d) Simulate.

Solution:

```
import matplotlib.pyplot as plt
from collections import Counter
from random import randint

def catch_them_all(n: int = 50) -> int:
    coupons = [False] * n
    coupons_collected = 0
    coupons_bought = 0
    while coupons_collected < len(coupons):
        new_coupon = randint(0, len(coupons) - 1)
```

```

    coupons_bought += 1
    if not coupons[new_coupon]:
        coupons[new_coupon] = True
        coupons_collected += 1
    return coupons_bought

cnt = Counter(catch_them_all(50) for _ in range(10000))
plt.bar(cnt.keys(), cnt.values())
plt.xlabel("Steps untill collecting all 50 coupons")
plt.ylabel("How many times did we take this many steps")
# plt.show()
plt.savefig('coupon_collector.pdf')

```

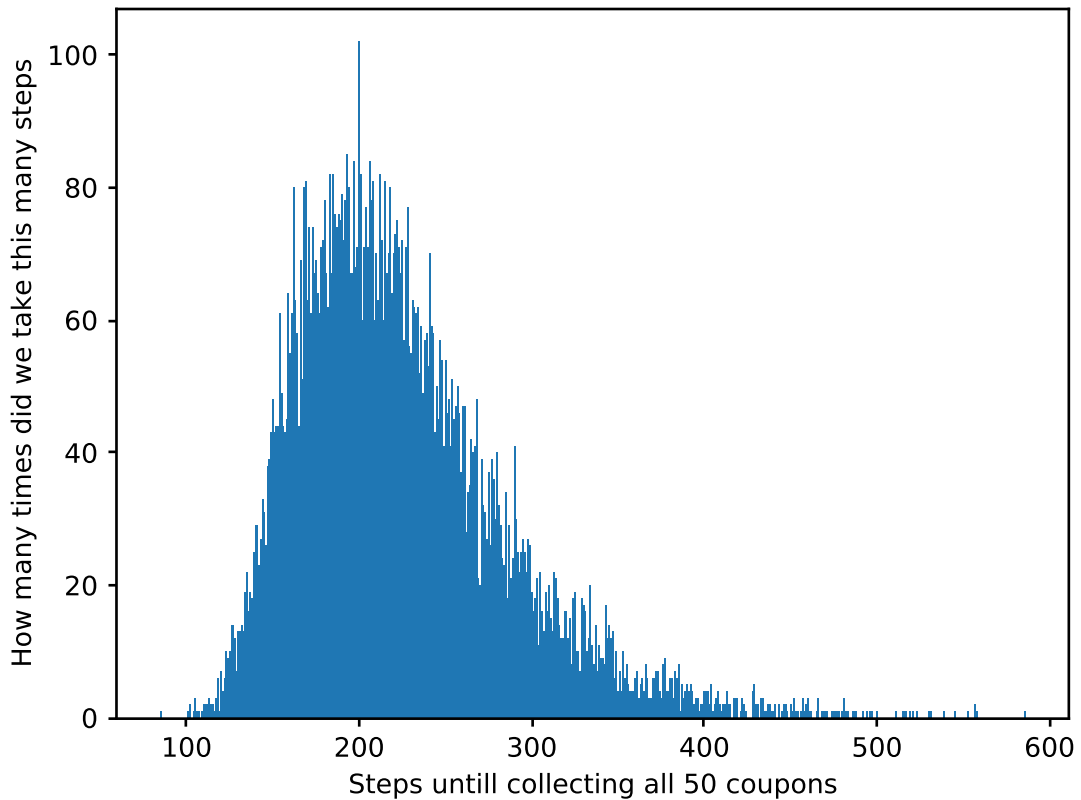


Figure 3.1: A histogram of how many steps were necessary (say 200 steps was necessary around 80 times).

$$h_{u,v} = \mathbb{E}[\text{\# of steps from } u \text{ to } v]$$

1) undirected

2) \leftarrow all edges were equally probable



3.2 Tutorial

1. Find a family of oriented graphs of constant in-degree and constant out-degree and as large hitting time as possible.

Solution: Let us first do constant out-degree and unbounded in-degree. We will later use a tree to achieve constant in-degree.

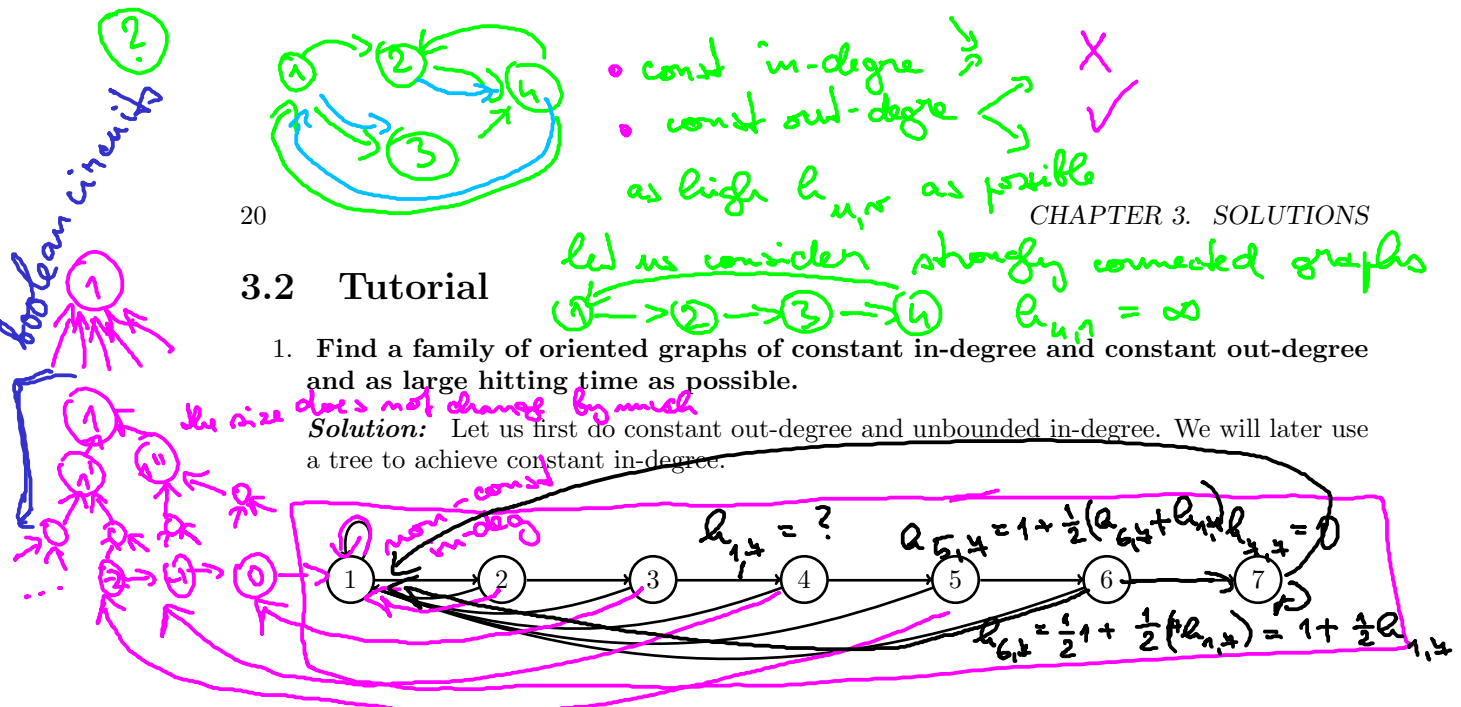


Figure 3.2: Oriented path with backwards arcs (oriented edges).

The expected hitting time from

$$h_{n,n} = 0$$

$$h_{n-1,n} = 1 + 0.5h_{1,n}$$

$$h_{n-2,n} = 1 + 0.5(h_{1,n} + h_{n-1,n})$$

$$= 1 + 0.5(h_{1,n} + (1 + 0.5h_{1,n}))$$

$$= 1.5 + 0.75h_{1,n}$$

$$h_{n-3,n} = 1 + 0.5(h_{1,n} + (1 + 0.5(h_{1,n} + (1 + 0.5h_{1,n}))))$$

$$h_{n-k,n} = \left(\sum_{j=0}^{k-1} 0.5^j \right) + \left(h_{1,n} \sum_{j=1}^k 0.5^j \right)$$

$$= 2 - 2^{1-k} + h_{1,n}(1 - 2^{-k})$$

Thus in particular when $k = n - 1$:

$$h_{1,n} = 2 - 2^{1-(n-1)} + h_{1,n}(1 - 2^{-(n-1)})$$

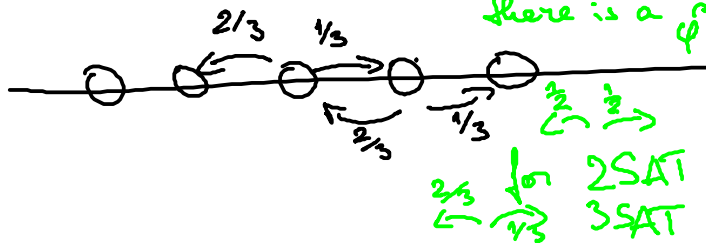
$$2^{-(n-1)}h_{1,n} = 2 - 2^{1-(n-1)}$$

$$h_{1,n} = 2^{n-1}(2 - 2^{1-(n-1)})$$

$$= 2^n - 2$$

exponential in # of vertices

Note that similar situation could happen on undirected graphs where the probabilities of traversing edge one way and the other way would not be the same. Which is in principle almost an oriented graph.



to the coloring alg from SAT

there is a "similar idea" alg. for SAT ϕ , start with rand. solution & change variables until $\phi(x) = 1$ poly $O(n^2)$ (we know deterministic $O(2^{2kn})$)

$$A + \begin{pmatrix} d & 0 \\ 0 & d \end{pmatrix}$$

M.C.

$$\pi_0, \pi_1 = A\pi_0$$

$$\pi_j = A^j \pi_0$$

3.2. TUTORIAL 2.

$\vec{0}$ is not an eigenvector!

2. Let $A \in \mathbb{R}^{n \times n}$ be a matrix with eigenvalues $\lambda_1, \dots, \lambda_n$. Show that the matrix $A + dI_n$ has eigenvalues $d + \lambda_1, \dots, d + \lambda_n$.

Solution: Eigenvalues and eigenvectors recap:

- We are interested in the limit of a Markov chain. When π_0 is the initial distribution, then $P^n \pi_0$ is the distribution after n steps.
- When we are iteratively multiplying a vector by a matrix from left, the simplest form we can hope for are eigenvectors, which satisfy

$$Ax = \lambda x$$

$$Ax - \lambda x = \vec{0}$$

$$Ax - \lambda I x = \vec{0}$$

$$(A - \lambda I)x = \vec{0} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$Ax = \lambda x \quad A^j x = \lambda^j x$$

Where A is a square matrix, λ is a real number called the *eigenvalue*, x is called the *eigenvector*. Then

$$A^n x = A(A^{n-1}x) = \lambda^n x$$

has nontrivial solution $\Leftrightarrow A - \lambda I$ is singular $\det(B) = 0 \Leftrightarrow B$ is singular

- For small matrices we usually use the characteristic polynomial:

$$\det(A - \lambda I) = 0$$

\hookrightarrow a polynomial in $\lambda \in \mathbb{C}[\lambda]$

the roots of this polynomial are the eigenvalues and we find the corresponding eigenvectors as:

$$A - \lambda I = \vec{0}$$

- For an eigenvalue λ we define its *algebraic multiplicity* to be the multiplicity of λ as the root of the characteristic polynomial. $(1-\lambda)(1-\lambda) = 0 \quad \lambda=1$ has alg.m.=2
- For an eigenvalue λ we define its *geometric multiplicity* to be the dimension of

$$B \in \mathbb{R}^{m \times m}$$

$$\text{Ker } B = \{x \in \mathbb{R}^m \mid Bx = \vec{0}\}$$

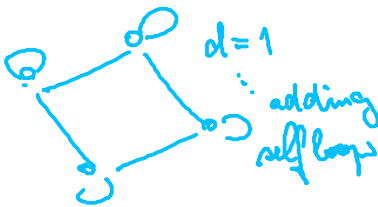
$$\dim(\text{Ker}(A - \lambda I))$$

$$1 \leq \text{geom.} \leq \text{alg.}$$

- We know that for any eigenvalue algebraic multiplicity is at least the geometric multiplicity.
- For each eigenvalue there is at least one eigenvector.
- It is usually infeasible to find roots of the characteristic polynomial when the matrix A is large. There are however computationally efficient methods of computing eigenvalues and eigenvectors (usually iterative multiplication converges to the eigenvector).

$$A^m x \rightarrow x_1$$

We use the definition, let λ be an eigenvalue in question and x its corresponding eigenvector:



$$\begin{aligned} Ax &= \lambda x \\ (A + dI)x &= Ax + dIx \\ &= \lambda x + dx \\ &= (\lambda + d)x \end{aligned}$$

Note that this can be rather useful when computing eigenvalues of a given matrix.

$$\|x\| = \sqrt{\langle x | x \rangle} = \sqrt{x^T x}$$

$A \in \mathbb{R}^{m \times m}$ $A = A^T$ $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be its eigenvalues.

$\lambda_1 \leq \max_{x, \|x\|=1} x^T A x$ x_1 $x_1^T A x_1 = x_1^T \lambda_1 x_1 = \lambda_1 x_1^T x_1 = \lambda_1$

(Rayleigh principle)

3. Show Courant-Fisher: Let $A \in \mathbb{R}^{n \times n}$ be a symmetric matrix ($A^T = A$). Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be its eigenvalues. Show

(a) $\lambda_1 = \max_{x \in \mathbb{R}^n, \|x\|=1} x^T A x$ $A^T = A$

Solution: Since A is Hermitian, we know that it is diagonalizable and we can choose an orthonormal basis of eigenvectors u_1, u_2, \dots, u_n . That is for any j we have $u_j^T u_j = 1$ and $Au_j = \lambda_j u_j$, and for any $i \neq j$ we have $u_j^T u_i = 0$.

We show two inequalities:

$$\max_{x \in \mathbb{R}^n, \|x\|=1} x^T A x \geq u_j^T A u_j = u_j^T \lambda_j u_j = \lambda_j$$

could happen $\alpha_1 = -5 \rightarrow \alpha_1^2 = 25$?

On the other hand we may write $x = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n$ and thus get

$$\begin{aligned} \max_{x \in \mathbb{R}^n, \|x\|=1} x^T A x &= (\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n)^T (A(\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n)) \text{ associativity} \\ &= (\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n)^T (\lambda_1 \alpha_1 u_1 + \lambda_2 \alpha_2 u_2 + \dots + \lambda_n \alpha_n u_n) \\ &= \lambda_1 \alpha_1^2 + \lambda_2 \alpha_2^2 + \dots + \lambda_n \alpha_n^2 \text{ (since } u_j^T u_i = 0 \text{ and } u_j^T u_j = 1) \\ &\leq \lambda_1 \end{aligned}$$

Where the last equation follows from the fact that if Q is orthogonal matrix ($Q^T Q = I$) then $\|x\| = \|Qx\|$ since $\|x\| = \sqrt{x^T x}$ and $\|Qx\| = \sqrt{x^T Q^T Q x}$.

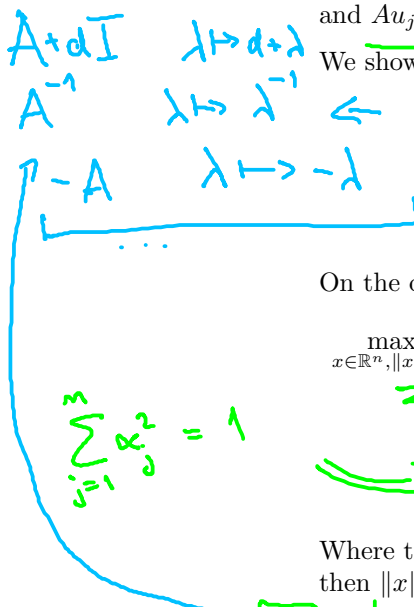
(b) $\lambda_n = \min_{x \in \mathbb{R}^n, \|x\|=1} x^T A x$

Solution: Consider $-A$ and use the previous result.

(c) The eigenvalue λ_2 can be computed similarly $\lambda_2 = \max_{x \in \mathbb{R}^n, \|x\|=1, x^T u_1 = 0} x^T A x$ (where u_1 is the eigenvector corresponding to λ_1). We can get other eigenvalues in a similar manner. Moreover we could use this to prove the interlacing theorem. See https://en.wikipedia.org/wiki/Min-max_theorem

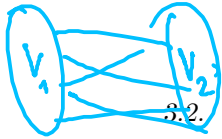
$\lambda_2 = \min_{\dots} \max_{\dots} x^T A x$ $\lambda_{n-2} = \max_{\dots} \min_{\dots} x^T A x$

if $\dim U + \dim V > n \Rightarrow U \cap V = \emptyset$
 $U, V \subseteq \mathbb{R}^n$

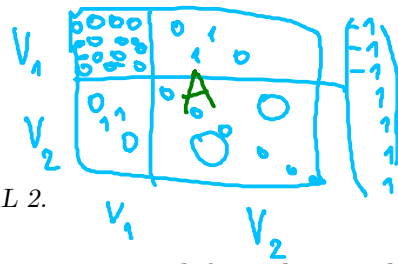


consider if $-d$ is eigenval. consider by the eig.v.
 $M = \max_{v \in V(G)} |x_v|$
 $A \mapsto dA$
 $\lambda \mapsto d\lambda$
 $(Ax)_v = \sum_{w \in N(v)} A_{vw} x_w$
 $Ax = d \cdot x$ neighbors of v
 $(dA)x = dAx$

G is bipartite



TUTORIAL 2.



$$A_{u,v} = \begin{cases} 1 & \text{if } (u,v) \in E \\ 0 & \text{otherwise} \end{cases}$$

$$A_{u,v} = \begin{cases} 1 & \text{if } (u,v) \in E \\ 0 & \text{otherwise} \end{cases}$$

4. Show that a connected d -regular graph is bipartite iff the least eigenvalue of its adjacency matrix is $-d$.

Solution: We know that the largest eigenvalue of the adjacency matrix of a d -regular graph is d and there is a corresponding eigenvector $(1, 1, \dots, 1)^T$.

If the graph is bipartite (that is $V(G) = A \cup B$ and $E(G) \subseteq A \times B$), we may use the vector defined as follows:

$$x_v = \begin{cases} -1 & \text{if } v \in A \\ 1 & \text{if } v \in B \end{cases}$$

Then x is an eigenvector corresponding to $-d$.

Let $(x_1, x_2, \dots, x_n)^T$ be the eigenvector corresponding to $-d$. Thus

$$-dx_i = \sum_{j \in N(i)} x_j$$

Let $M = \max_i |x_i|$ and $P = \{i \mid x_i = M\}$ and $N = \{i \mid x_i = -M\}$. Without loss of generality let P be non-empty. For any $i \in P$ we have

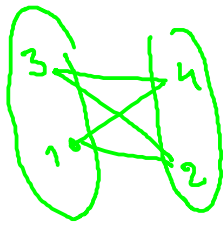
$$-dM = \sum_{j \in N(i)} x_j$$

$\forall_j |x_j| \leq M$
 $\Rightarrow \forall_j \in N(i) \quad x_j = -M$

thus $x_j = -M$ for each $j \in N(i)$ (since each $|x_k| \leq M$).

Since the graph is connected we eventually get that for any i it holds that $|x_i| = M$.

Eigenvalues of the graph of neurons in human brain have been considered in epilepsy – they studied “how much” is the brain bipartite, which can be expressed by the difference between the smallest eigenvalue and the negative degree.



$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = -d \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

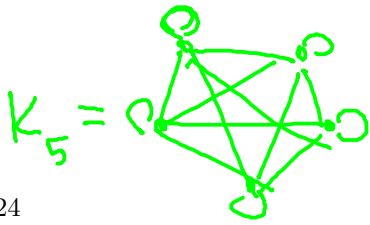
$$-2M = x_2 + x_4 \quad |x_j| \leq M$$

\uparrow
 $d=2$

eigenvalues $\leq \frac{n+1}{2}$
 eigenv. $A \dots \lambda$
 eigenv. $A + I \dots \lambda + 1$
 $d=1$

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

$n-1$ alg. m. 1
 geom. m. 1
 -1 alg. m. $n-1$
 geom. m. $= n-1$



$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

5. Compute the eigenvalues and eigenvectors of the following graphs:

(a) K_n , the complete graph on n vertices.

Solution: It will be easier to determine eigenvalues and eigenvectors of a complete graph with selfloops (we add unit matrix). We may subtract ones if we mind the selfloops.

The adjacency matrix of a complete graph with selfloops is:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad A^T = A$$

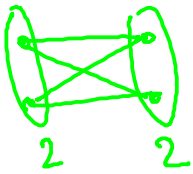
By the observation from the lecture we know that this is a regular graph and the matrix above has eigenvalue n with eigenvector $(1, 1, 1, 1, 1)^T$. By Hamiltonicity we know that all other eigenvectors are perpendicular to the one above. Thus all their entries sum up to zero.

We guess other eigenvectors (we need to guess $n - 1$ of them).

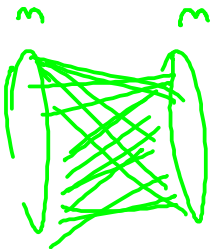
$(1 \rightarrow 000)$ $\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

$$= 1 \cdot 1 + (-1) \cdot 1 + 0 \cdot 1 + 0 \cdot 1 + 0 \cdot 1 = 0$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$



$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$



n -regular
eigenval n

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

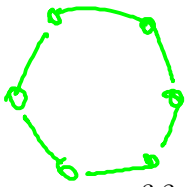
$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Thus geometric (and thus also the algebraic) multiplicity of eigenvalue 0 is $n - 1$.

(b) $K_{n,n}$, the complete bipartite graph with partites of size n each.

Solution: Here we are happy with no selfloops (otherwise the graph would not even be bipartite).

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$



eigenw. -- \cos

3.2. TUTORIAL 2.

The graph is bipartite and regular, thus we know that the largest eigenvalue is n with the eigenvector $(1, 1, 1, 1, 1, 1)^T$ the smallest eigenvalue is $-n$ with the eigenvector $(-1, -1, -1, 1, 1, 1)^T$. As with the complete graph with selfloops it is easy to show that the rest is zero eigenvalues with corresponding vectors.

(c) C_n , the cycle on n vertices.

Solution: If we knew circular matrices we could use their properties. We will write the adjacency matrix as a sum of two simpler matrices:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Observe that moreover the two matrices are inverse to each other, thus their eigenvalues are inverses as:

$$\begin{aligned} Ax &= \lambda x \\ x &= Ax \\ &= A^{-1}Ax \\ &= \lambda A^{-1}x \end{aligned}$$

Let $\omega \in \mathbb{C}$ be the primitive n -th root of unity. Thus $\omega = e^{2i\pi/n}$.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \omega^0 \\ \omega^1 \\ \omega^2 \\ \omega^3 \\ \omega^4 \\ \omega^5 \end{pmatrix} = \begin{pmatrix} \omega^5 \\ \omega^0 \\ \omega^1 \\ \omega^2 \\ \omega^3 \\ \omega^4 \end{pmatrix} = \omega^5 \begin{pmatrix} \omega^0 \\ \omega^1 \\ \omega^2 \\ \omega^3 \\ \omega^4 \\ \omega^5 \end{pmatrix}$$

Similarly for the even powers

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \omega^0 \\ \omega^2 \\ \omega^4 \\ \omega^0 \\ \omega^2 \\ \omega^4 \end{pmatrix} = \begin{pmatrix} \omega^4 \\ \omega^0 \\ \omega^2 \\ \omega^4 \\ \omega^0 \\ \omega^2 \end{pmatrix} = \omega^4 \begin{pmatrix} \omega^0 \\ \omega^2 \\ \omega^4 \\ \omega^0 \\ \omega^2 \\ \omega^4 \end{pmatrix}$$

And so on. A particular eigenvector is an eigenvector of the eigenvalue ω^j with respect to this matrix and of eigenvalue ω^{-j} with respect to the inverse matrix. Thus when we sum the two matrices we get that the eigenvalue is $\omega^j + \omega^{-j}$. Observe that $\omega^j + \omega^{-j} \in \mathbb{R}$.

6. Distributed discrete logarithm algorithm.

Solution: We say that a group \mathbb{G} is *cyclic* iff any of its element can be generated using a single generator. Say we have the group (\mathbb{Z}_5^*, \cdot) with its generator 3:

$$\mathbb{Z}_5^* = \{1, 2, 3, 4\}$$

$$3^0 = 1$$

$$3^1 = 3$$

$$3^2 = 4$$

$$3^3 = 2$$

When we fix a group \mathbb{G} and its generator g we may ask what is the discrete logarithm of a given group element:

$$DLog_{\mathbb{G},g}(x) = \min_{n \in \mathbb{N}} g^n = x$$

So for instance $DLog_{\mathbb{Z}_5^*,3}(4) = 2$.