به نام خدا

دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

درس شناسایی آماری الگو

استاد رحمتی

تمرین اول

علیرضا مازوچی

۴۰۰۱۳۱۰۷۵

سوال ۱

پیش از جواب‌دادن به سوالات لازم است تا روش طراحی سامانه را تبیین کنم. من قصد دارم تا سه دسته‌بند طراحی کنم و سپس این سه دسته‌بند در یک چهارچوب گروهی (Ensemble) جواب نهایی را برگردانند. دسته‌بند اول با چهره شخص و یک شبکه عصبی تشخیص می‌دهد که فرد با چه احتمالی در بازه مناسب قرار دارد. دسته‌بند دوم با پردازش صوت کودک با یک شبکه عصبی دیگر تشخیص می‌دهد که کودک به چه احتمالی در بازه سنی مناسب قرار دارد یا نه؟ دسته‌بند سوم با ویژگی قد و وزن کودک احتمال تعلق داشتن کودک به بازه سنی مناسب را بیان می‌کند. نهایتا خروجی هر سه دسته‌بند با به یک MLP ساده می‌دهیم تا نتیجه نهایی را اعلام کند.

الف) مسئله از نوع دسته‌بندی یا Classification است؛ چراکه قرار است کل بازدیدکنندگان را به دو کلاس گروه سنی مناسب و گروه سنی نامناسب تقسیم کنیم.

ب) به سنسور تشخیص چهره، صوت، قد و وزن نیاز داریم.

ج) ما یک مجموعه‌داده نیاز داریم که برای هر داده یک تصویر از تمام‌رخ چهره، یک صوت بعد از خواندن یک متن، قد و وزن را در خود داشته باشد.

د) می‌توان داده‌ها را از بازدیدکنندگان فعلی تا زمانی استقرار سامانه تهیه کرد. یعنی هر تصویر و سایر مشخصات هر کودک بازدیدکننده را می‌توان برداشت و یک ناظر انسانی مسئول تعیین کلاس کودک می‌شود. باتوجه به آمار بالای بازدید پارک دیزنی، با این شیوه می‌توان در مدت زمان قابل قبولی حجم مناسب داده جمع‌آوری شود. البته باید توجه داشت ممکن است قوانین حفظ حریم خصوصی مانع از جمع‌آوری داده تمام بازدیدکننده‌ها بدون اجازه آن‌ها باشد که در این صورت لازم است این جمع‌آوری داوطلبانه و با آگاهی بازدیدکننده‌ها صورت بگیرد.

علاوه بر شیوه فوق می‌توان از مجموعه‌داده‌های آماده هم کمک گرفت که البته هر کدام بخشی از ویژگی‌ها مدنظر ما را دارند و می‌توانند در توسعه زیرمدل‌های ما مورد استفاده قرار بگیرند.

ه) ویژگی قد، وزن، یک تصویر که یک آرایه دوبعدی از ویژگی‌هاست و صوت که یک آرایه یک‌بعدی از ویژگی‌هاست.

و) برای ویژگی قد و وزن پیش‌پردازش خاصی نیاز نیست. در مورد صوت، لازم است تا پیش‌پردازشی برای حذف نویز محیط و سایر نویزهای این حوزه انجام گیرد. در مورد تصویر هم باید پیش‌پردازش‌هایی برای شفاف کردن عکس، حذف لرزش موقع تصویر و تاری عکس و مواردی از این دست صورت بگیرد.

ز) مهم‌ترین چالش جمع‌آوری داده است. باتوجه به پیچیدگی مدل و دارابودن انواع ویژگی‌ها نیاز به تعداد داده بالاست. قوانین حفظ حریم خصوصی هم ممکن است کار را سخت‌تر بکنند. در این حالت ممکن است داده به اندازه جمع‌آوری نشود که مجبور خواهیم بود از مدل‌های ساده برای بخش تصویر و صوت استفاده کنیم که بی‌شک در دقت مدل موثر است. حتی این احتمال وجود دارد که مجبور باشیم بخشی از مدل را مثلا بخش تصویر را حذف کنیم.

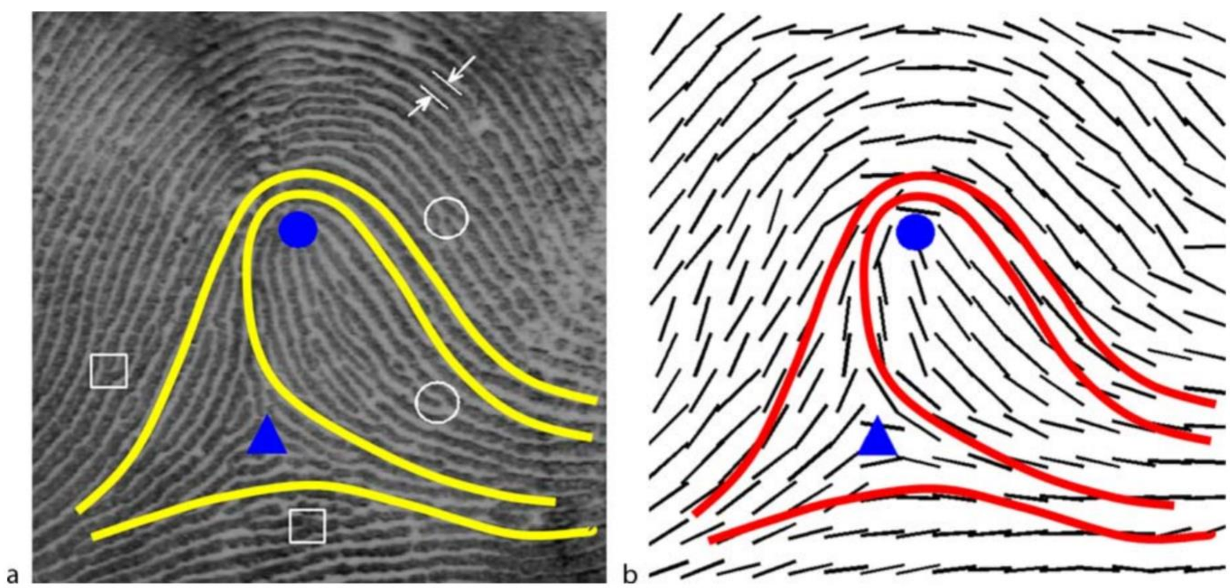
یکی دیگر از چالش‌های سیستم جدید، امکان تغلب در آن است. درحالی که ویژگی‌هایی مانند قد به سختی مورد تغلب واقع می‌شود ولی در مورد صوت و تصویر ممکن است کاربر تغییراتی در صدا و چهره خود بدهد که سیستم را به شک بیاندازد. در مورد وزن هم امکان افزایش آن با لباس‌های سنگین وجود دارد. همچنین به دلیل کرونا (و بیماری‌های آتی!) ممکن است امکان دریافت چهره کامل به سختی صورت بگیرد.

ح) نقطه قوت اصلی سیستم، دقت بیشتر و توجه به جوانب مختلف نسبت به سیستم قبلی است. در سیستم قبلی تنها به قد توجه می‌شد که ممکن بود یک کودک که قد آن نسبت سن واقعی‌اش متداول نباشد را به خطا بیاندازد درحالی که در این سیستم مبتنی بر چندین زیرمدل است. نقطه قوت دیگری که شاید بتوان برای آن درنظر گرفت آن است که خروجی سیستم فعلی برای کودکان و والدینشان پذیرفتنی‌تر است. در سیستم قبلی محتمل بود تا برخی از کودکان و والدین توجه تنها به یک خط‌کش را عادلانه ندانند درحالی که در سیستم فعلی و پس از جمع‌آوری اطلاعات مختلف و نهایتا تصمیم‌گیری توسط کامپیوتر امکان متقاعدکردن بهتری وجود خواهد داشت.

درکنار نقاط قوت، نقاط ضعفی وجود دارد. سیستم فعلی نسبت به سیستم قبل، پیچده‌تر است و برای استفاده از آن هم نیاز به تجهیزات مختلف سخت‌افزاری و نرم‌افزاری است. استفاده از سیستم پیشنهادی وقت کاربران را می‌گیرد چراکه آن‌ها باید چهار تست را به جای یک تست بدهند. اگرچه می‌توان این‌ها را باهم گرفت. (مثلا در حین ایستادن بر روی وزنه، قد و تصویر هم برداشته شود و در همین حین از کودک خواسته شود یک متن یا شعر را بخواند.) همچنین جمع‌آوری داده‌ها همانطور که گفته شد سخت، زمان‌بر و پرهزینه است.

سوال ۲

الف) می‌توان محل تورفتگی‌ها را مانند یک خط به یکدیگر وصل کرد. سپس از الگوهای این خط‌ها ویژگی‌هایی را پیشنهاد داد. مثلا می‌توان محلی که این خط‌ها سه‌شاخه شده است را ذخیره کرد؛ مانند مربع‌های توخالی در شکل زیر. همچنین می‌توان محل‌هایی که خط‌ها خاتمه پیدا کردند را هم به عنوان ویژگی ذخیره کرد؛ مانند دایره‌های توخالی در شکل زیر.[1]



[1] https://sci-hub.se/10.1007/978-0-387-73003-5_50

ب) وضعیت لب‌ها و وضعیت ابرو می‌تواند دو ویژگی قابل محاسبه و مفید برای این مسئله باشد. توجه کنید که در این حالت لب‌ها و ابروها را مانند قسمت قبل با یک خط مدل می‌کنیم. این ویژگی‌های ساده‌ای چراکه به آسانی می‌توان دید در احساسات مختلف وضعیت لب و ابرو متفاوت است.

ج) می‌توان تصاویر متوالی از یک فرد را دریافت کرد و میزان حرکت سر و دست او را به عنوان ویژگی حساب کرد. (با اختلاف تصاویر) چراکه برخی از افراد هنگام حرکت، دست‌هایشان یا سرشان بیشتر تحرک دارد. همچنین می‌توان نسبت اندازه قسمت‌های مختلف بدن به یکدیگر در یک تصویر را هم درنظر گرفت.[۲]

د) با توجه به داشتن تصاویر متوالی از یک فرد می‌توان ویژگی توقف را برای هر شخص تعریف کرد. این ویژگی برای تشخیص فرد ایستاده متوقف، نشسته و مانند آن مناسب است. همچنین می‌توان پس از تشخیص هر فرد، نسبت ارتفاع به عرض کوچکترین مستطیلی که شخص در آن قرار گرفته است را هم بررسی کرد. این ویژگی هم کمک می‌کند متوجه شویم فرد روی پا ایستاده یا خیر.

ه) تعداد گوشه (تعداد شکستگی)، نسبت پیکسل‌های الگو در حدود بالاترین (یا پایین‌ترین، یا راست‌ترین یا چپ‌ترین) قسمت الگو نسبت به کل پیکسل‌ها، نسبت مساحت محصور در الگو به محیط الگو

و) تعداد گوشه (تعداد شکستگی)، نسبت مجموع پیکسل‌های الگو در حدود بالاترین و پایین‌ترین (یا مجموع راست‌ترین و چپ‌ترین) قسمت الگو نسبت به کل پیکسل‌ها، نسبت مساحت محصور در الگو به محیط الگو

ز) تعداد گوشه (تعداد شکستگی)

ح) تعداد گوشه (تعداد شکستگی)، نسبت پیکسل‌های الگو در حدود بالاترین (یا پایین‌ترین، یا راست‌ترین یا چپ‌ترین) قسمت الگو نسبت به کل پیکسل‌های الگو، نسبت مساحت محصور در الگو به محیط الگو

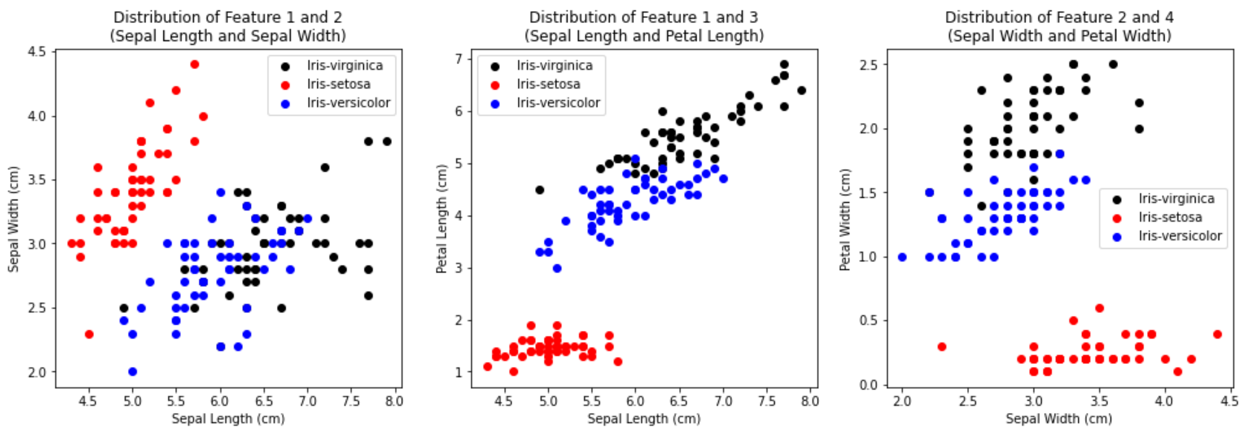
[2] https://sci-hub.se/10.1117/12.2015634

ط) تعداد گوشه (تعداد شکستگی)، نسبت مساحت محصور در الگو به محیط الگو

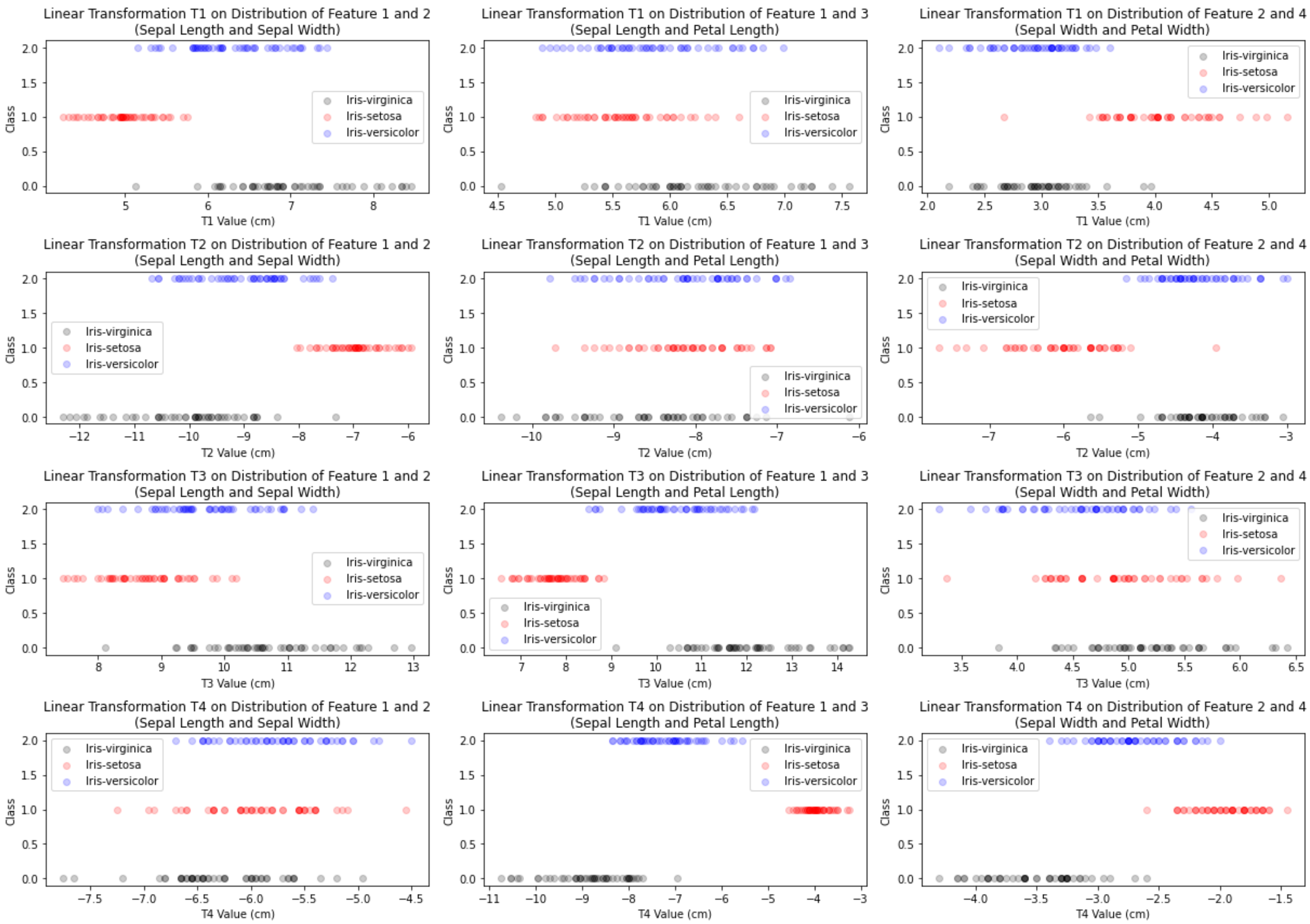
ی) تعداد گوشه (تعداد شکستگی)، نسبت مساحت محصور در الگو به محیط الگو، مجموع پیکسل‌ها در بالاترین، پایین‌ترین، راست‌ترین و چپ‌ترین قسمت الگو نسبت به کل پیکسل‌های الگو (به شرط عدم چرخش مربع و مثلث)

سوال ۳

الف)



ب)

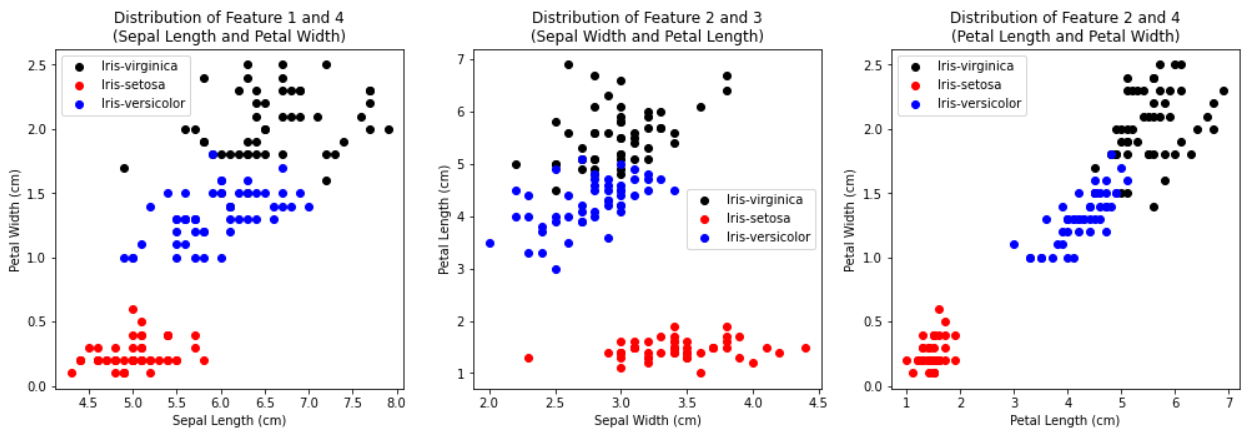


ج) برای ویژگی ۱ و ۲، T_2 از همه مناسب‌تر به نظر می‌رسد اگرچه T_1 هم قابل قبول است. دو تبدیل T_3 و T_4 نتوانسته‌اند برخلاف دو تبدیل دیگر، کلاس نسبتا مجزای قرمز را به خوبی جداکنند. در مورد T_2 و T_2 هر دو تقریبا مشابه هستند اما تبدیل T_2 کمی کلاس قرمز را بهتر جدا کرده است.

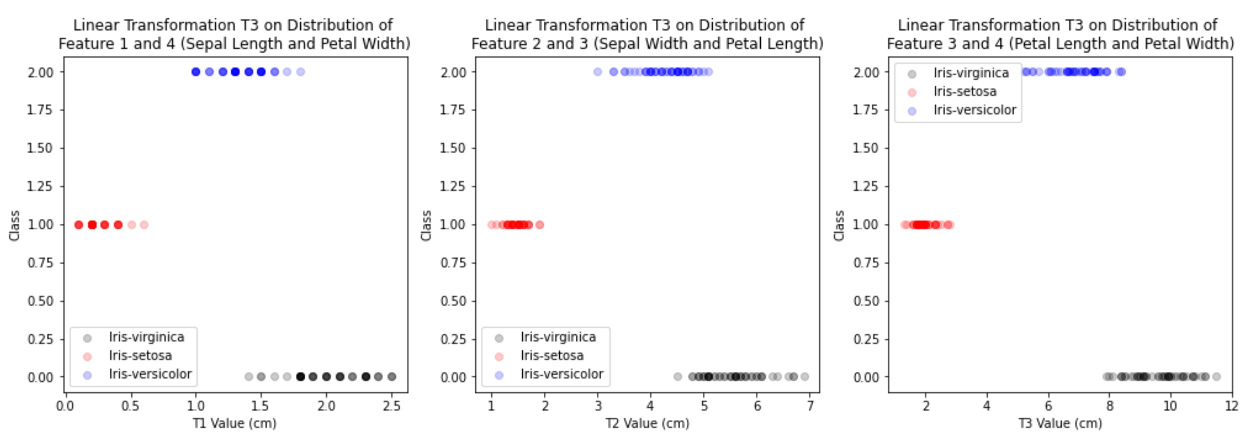
برای ویژگی ۱ و ۳، T_4 مناسب‌ترین است. در تبدیل T_1 و T_2 داده‌های هر سه کلاس روی هم افتاده است. در تبدیل T_3 هم دو کلاس آبی و سیاه همپوشانی زیادی دارند و داده‌های کلاس قرمز نسبت به T_4 پراکنده‌تر است.

برای ویژگی ۲ و ۴، T_4 مناسب‌ترین است. تبدیل T_3 داده‌های هر سه کلاس را در یک نقطه نگاشت می‌کند. T_1 و T_2 دو کلاس آبی و سیاه را روی هم می‌اندازد درحالی که در T_4 سه کلاس تاحد قابل قبولی از هم جدا شده‌اند.

د)



ه) با توجه به نمودارهای قسمت د می‌توان حدس زد که برای جفت ویژگی (۱ و ۴) و (۲ و ۳) خطوط موازی با محور xها با دقت بسیار خوبی می‌توانند داده‌های کلاس‌های مختلف را از هم جدا کنند. لذا برای این دو قسمت از تبدیل خطی [۱ و ۰] استفاده کردم. برای جفت ویژگی (۲ و ۴) شـــاید خطوط مورب ۴۵ درجه (بدون توجه به مقیاس) به یکی از بهترین نتایج ممکن ختم شـــود. نهایتا با اختلاف مقیاس دوبرابر ویژگی اول و دوم، از تبدیل خطی [۲ و ۱] استفاده کردم:



سوال ۴

الف) پرتاب سکه از توزیع دوجمله‌ای پیروی می‌کند.

$$q = p(tail) = \frac{2}{3} \to p = p(head) = 1 - q = 1 - \frac{2}{3} = \frac{1}{3}$$

$$\bar{X} = n * p = 1400 * \frac{1}{3} = \frac{1400}{3} \approx 466.67$$

$$\sigma_X^2 = n * p * q = 1400 * \frac{1}{3} * \frac{2}{3} = \frac{2800}{9}$$

$$\sigma_x = \sqrt{\sigma_X^2} = \sqrt{\frac{2800}{9}} = \frac{20}{3}\sqrt{7} \approx 17.64$$

ب) برای قسمت اول این سوال باید سه توپ از شش توپ را انتخاب کرد به عنوان شوت موفق و سپس باید سه توپ با احتمال موفقیت گل شوند و سه توپ دیگر با احتمال شکست به بیرون بروند. برای قستم دوم هم باید شوت اول و دوم بیرون برود و شوت سوم حتما باید گل شود. مابقی شوت‌ها اهمیتی ندارد.

$$p = p(success) = \frac{4}{10} = \frac{2}{5} \to q = p(fail) = 1 - p = \frac{3}{5}$$

$$p\ (hat - trick) = \binom{6}{3} * p^3 * q^{(6-3)} = \binom{6}{3} * \left(\frac{2}{5}\right)^3 * \left(\frac{3}{5}\right)^3 = 20 * \left(\frac{6}{25}\right)^3 = 0.27648$$

$$p(f, f, s, *) = q * q * p = \frac{3}{5} * \frac{3}{5} * \frac{2}{5} = 0.144$$

ج)

$$P\ (X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$P(X = 3) = P(X = 0) + P\ (X = 1) \to \frac{\lambda^3 e^{-\lambda}}{3!} = \frac{\lambda^0 e^{-\lambda}}{0!} + \frac{\lambda^1 e^{-\lambda}}{1!} \to \frac{\lambda^3 e^{-\lambda}}{6} = (1 + \lambda)e^{-\lambda}$$

$$\to \frac{\lambda^3}{6} = 1 + \lambda \to \lambda^3 - 6\lambda - 6 = 0$$

این معادله را من به صورت دستی نتوانستم حل کنم. با سپردن این معادله به ابزارهای آنلاین متوجه شدم این معادله تنها یک جواب حقیقی دارد و آن در حدود ۲/۸۴۷ بوده است.

$$\bar{X} = \lambda \approx 2.847$$

برای قسمت بعدی سوال از جدول مقادیر تجمعی این توزیع[3] و $\lambda = 2.8$ استفاده کردم.

$$P(2 \le X \le 4) = P(X \le 4) - P(x \le 2) = 0.848 - 0.469 = 0.379$$

د)

$$E[aX + bY] = \frac{1}{n}\sum_i (ax_i + by_i) = \frac{1}{n}\sum_i ax_i + \frac{1}{n}\sum_i by_i = \frac{a}{n}\sum_i x_i + \frac{b}{n}\sum_i y_i$$
$$= a * \left(\frac{1}{n}\sum_i x_i\right) + b * \left(\frac{1}{n}\sum_i y_i\right) = aE[x] + bE[y]$$

$$\sigma_X^2 = E[(X - \mu_X)^2] = E[X^2 - 2X\mu_X + \mu_X^2] = E[X^2] - 2E[X\mu_X] + E[\mu_X^2]$$
$$= E[X^2] - 2\mu_X E[X] + \mu_X^2 = E[X^2] - 2\mu_x^2 + \mu_X^2 = E[X^2] - \mu_x^2$$

$$X \text{ and } Y \text{ are independent} \to E[XY] = E[x]E[y]$$

$$\rho(X,Y) = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} = \frac{E[XY - X\mu_Y - Y\mu_X + \mu_X \mu_Y]}{\sigma_X \sigma_Y}$$
$$= \frac{E[XY] - E[X\mu_Y] - E[Y\mu_X] + E[\mu_X \mu_Y]}{\sigma_X \sigma_Y}$$
$$= \frac{E[X]E[Y] - \mu_Y E[x] - \mu_X E[Y] + \mu_X \mu_Y}{\sigma_X\ \sigma_Y} = \frac{\mu_X \mu_Y - \mu_x \mu_Y - \mu_x \mu_Y + \mu_x \mu_Y}{\sigma_X \sigma_Y} = 0$$
$$\to X \text{ and } Y \text{ are uncorralated.}$$

برای قسمت چهارم این سوال می‌توانیم با یک مثال نقض نشان دهیم که ممکن است دو ویژگی X و Y همبستگی نداشته باشند ولی با یکدیگر رابطه داشته باشند.

$$Y = X^2$$

$$X = \{-2, -1, 0, +1, +2\}, Y = \{4, 1, 0, 1, 4\}$$

[3] https://www.york.ac.uk/depts/maths/tables/poisson.pdf

$$\mu_X = 0\,, \mu_Y = \frac{10}{3}$$

$$cov(X,Y) = \frac{1}{5}\left(-2 * \frac{2}{3} + 1 * \frac{7}{3} + 0 - 1 * \frac{7}{2} + 2 * \frac{2}{3}\right) = 0 \rightarrow \rho(X,Y) = 0$$

در این مثال X و Y با یکدیگر رابطه دارند ولی همبستگی‌شان برابر با صفر است!

سوال ۵

الف)

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, Av = \lambda v, v = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} a+b \\ c+d \end{bmatrix} = \begin{bmatrix} \lambda \\ \lambda \end{bmatrix} \rightarrow \lambda = a+b = c+d$$

طبق فرض سوال $a+b$ برابر با $c+d$ است که اثبات می‌کند v_1 یک بردار ویژه است و مقدار ویژه آن برابر با $a+b$ است.

ب)

$$a+b = c+d \rightarrow a-c = d-b$$

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, Av = \lambda v, v = \begin{bmatrix} b \\ -c \end{bmatrix} \rightarrow \begin{bmatrix} ab-bc \\ bc-cd \end{bmatrix} = \begin{bmatrix} \lambda b \\ \lambda c \end{bmatrix} \rightarrow \begin{bmatrix} b\,(a-c) \\ -c\,(\,d-b) \end{bmatrix} = \begin{bmatrix} \lambda b \\ -\lambda c \end{bmatrix}$$
$$\rightarrow \lambda = a-c = d-b$$

به طور مشابه در می‌یابیم v_2 نیز یک بردار ویژه است و مقدار آن برابر با $a-c$ یا $d-b$ است.

ج)

$$Av - \lambda v = 0 \rightarrow v(A - \lambda I) = 0$$

$$\rightarrow |A - \lambda I| = 0 \rightarrow \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = 0 \rightarrow (a-\lambda)(d-\lambda) - bc = 0$$

$$\rightarrow\ ad - \lambda(a+d) + \lambda^2 - bc = 0 \rightarrow \lambda^2 - \lambda(a+d) + (ad-bc) = 0$$

$$\rightarrow \lambda = \frac{(a+d) \pm \sqrt{(a+d)^2 - 4(ad-bc)}}{2} = \frac{(a+d) \pm \sqrt{a^2+d^2-2ad+4bc}}{2}$$

$$\rightarrow \lambda_1 = \frac{(a+d) + \sqrt{a^2+d^2-2ad+4bc}}{2}\ , \lambda_2 = \frac{(a+d) - \sqrt{a^2+d^2-2ad+4bc}}{2}$$

با این رابطه می‌توان دو مقدار ویژه را پیدا کرد.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, Av = \lambda v, v = \begin{bmatrix} e \\ f \end{bmatrix} \rightarrow Av - \lambda v = 0 \rightarrow (A - \lambda I)v = 0 \rightarrow \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} \begin{bmatrix} e \\ f \end{bmatrix} = 0$$

$$\rightarrow \begin{bmatrix} (a-\lambda)e + bf \\ ce - (d-\lambda)f \end{bmatrix} = 0 \rightarrow \begin{cases} (a-\lambda)e + bf = 0 \\ ce - (d-\lambda)f \ = 0 \end{cases} \xrightarrow{b \neq 0} f = \frac{\lambda - a}{b} e = \frac{c}{d-\lambda} e$$

در رابطه فوق برابر بودن $\frac{\lambda-a}{b}$ با $\frac{c}{d-\lambda}$ با استفاده از λ_1 و λ_2 برقرار خواهد بود و نهایتا متناظر با این دو مقدار ویژه، بردارهای ویژه زیر را خواهیم داشت:

$$v_1 = \begin{bmatrix} 1 \\ \frac{\lambda_1 - a}{b} \end{bmatrix}\ , v_2 = \begin{bmatrix} 1 \\ \frac{\lambda_2 - a}{b} \end{bmatrix}$$

این رابطه در شرایطی برقرار است که $b \neq 0$. حال اگر $b = 0$ باشد ولی $c \neq 0$ باشد می‌توان نوشت:

$$\begin{cases} (a-\lambda)e + bf = 0 \\ ce - (d-\lambda)f \ = 0 \end{cases} \xrightarrow{C \neq 0} e = \frac{d-\lambda}{c} f = \frac{b}{\lambda - a} f$$

$$\rightarrow\ v_1 = \begin{bmatrix} 1 \\ \frac{d-\lambda_1}{c} \end{bmatrix}\ , v_2 = \begin{bmatrix} 1 \\ \frac{d-\lambda_2}{c} \end{bmatrix}$$

اما اگر $b = c = 0$ باشد باید به نحو دیگری عمل کرد:

$$\begin{cases} (a-\lambda)e + bf = 0 \\ ce - (d-\lambda)f \ = 0 \end{cases} \xrightarrow{b=c=0} \begin{cases} (a-\lambda)e = 0 \\ (\lambda - d)f = 0 \end{cases}$$

$$\lambda = \frac{(a+d) \pm \sqrt{a^2+d^2-2ad+4bc}}{2} = \frac{(a+d) \pm \sqrt{a^2+d^2-2ad}}{2} = \frac{(a+d) \pm |a-d|}{2}$$

$$= \frac{(a+d) \pm (a-d)}{2} \rightarrow \lambda_1 = a\ , \lambda_2 = d$$

اگر $a \neq d$ باشد یک عبارت صفر می‌شود لذا بردارهای ویژه برابر خواهد بود با:

$$v_1=\begin{bmatrix}1\\0\end{bmatrix},v_2=\begin{bmatrix}0\\1\end{bmatrix}$$

د)

$$Av=\lambda v\to(A-\lambda I)v=0\to|A-\lambda I|=0\to\begin{vmatrix}0.75-\lambda & 0.25\\0.25 & 0.75-\lambda\end{vmatrix}=0$$
$$\to(0.75-\lambda)^2-0.25^2=0\to0.75-\lambda=\pm0.25\to\lambda_1=0.5,\lambda_2=1$$

$$\begin{bmatrix}0.25 & 0.25\\0.25 & 0.25\end{bmatrix}v=0\to\begin{bmatrix}0.25 & 0.25\\0.25 & 0.25\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=0\to\begin{bmatrix}0.25x+0.25y\\0.25x+0.25y\end{bmatrix}=0\to x=-y\to v_1=\begin{bmatrix}1\\-1\end{bmatrix}$$

$$\begin{bmatrix}-0.25 & 0.25\\0.25 & -0.25\end{bmatrix}v=0\to\begin{bmatrix}-0.25 & 0.25\\0.25 & -0.25\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=0\to\begin{bmatrix}-0.25x+0.25y\\0.25x-0.25y\end{bmatrix}=0\to x=y\to v_2=\begin{bmatrix}1\\1\end{bmatrix}$$

ه)

$$Ad_*=d_*,d_*=\begin{bmatrix}x\\y\end{bmatrix}\to\begin{bmatrix}0.75 & 0.25\\0.25 & 0.75\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}x\\y\end{bmatrix}\to\begin{bmatrix}0.75x+0.25y\\0.25x+0.75y\end{bmatrix}=\begin{bmatrix}x\\y\end{bmatrix}$$
$$\to\begin{cases}0.75x+0.25y=x\\0.25x+0.75y=y\end{cases}\to\begin{cases}-0.25x+0.25y=0\\0.25x+0.75y=0\end{cases}\to\begin{cases}x-y=0\\x+3y=0\end{cases}\to x=y=0$$
$$\to d_*=\begin{bmatrix}0\\0\end{bmatrix}$$

به نظر می‌رسد هیچ وضعیتی غیرصفری وجود ندارد که ثابت بماند.

و) اگر $d[0]=\begin{bmatrix}x\\-x\end{bmatrix}$ باشد لذا برای مقدار $d[1]$ خواهیم داشت (از روی مقدار و بردار ویژه هم می‌توانستیم این را بفهمیم):

$$d[1]=Ad[0]=\begin{bmatrix}0.75 & 0.25\\0.25 & 0.75\end{bmatrix}\begin{bmatrix}x\\-x\end{bmatrix}=\begin{bmatrix}0.5x\\-0.5x\end{bmatrix}=\frac{1}{2}d[0]$$

با توجه به اینکه $d[1]$ برابر ضریبی از $d[0]$ شد می‌توان نتیجه گرفت:

$$d[i]=\left(\frac{1}{2}\right)^i d[0]\to\lim_{n\to\infty}d[n]=\lim_{n\to\infty}\left(\frac{1}{2}\right)^n d[0]=\begin{bmatrix}0\\0\end{bmatrix}$$

لذا داستان خسرو و شیرین به این شکل خواهد شد که در ابتدا خسرو مقداری شیرین را دوست دارد و شیرین به همان اندازه از او متنفر است (یا بالعکس). سپس با گذشت هر لحظه این دوستی و تنفر به نصف مقدار قبل کاهش می‌یابد تا سرانجام هیچ حسی میان این دو نفر وجود نداشته باشد.

ز)

$$d[1] = Ad[0] = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 4.5 \end{bmatrix}$$

$$d[2] = Ad[1] = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \begin{bmatrix} 3.5 \\ 4.5 \end{bmatrix} = \begin{bmatrix} 3.75 \\ 4.25 \end{bmatrix}$$

$$d[3] = Ad[2] = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \begin{bmatrix} 3.75 \\ 4.25 \end{bmatrix} = \begin{bmatrix} 3.875 \\ 4.125 \end{bmatrix}$$

به نظر می‌رسد که هر دو بعد آرایه در حال نزدیک‌شدن به ۴ هستند. می‌توان این مورد را دقیق‌تر بررسی کرد:

$$d[0] = \begin{bmatrix} 4+1 \\ 4-1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$d[1] = Ad[0] = A\left(\begin{bmatrix} 4 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = A\begin{bmatrix} 4 \\ 4 \end{bmatrix} + A\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$d[i] = A^i d[0] = A^i\left(\begin{bmatrix} 4 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = A^i\begin{bmatrix} 4 \\ 4 \end{bmatrix} + A^i\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} + \left(\frac{1}{2}\right)^i\begin{bmatrix} 1 \\ -1 \end{bmatrix} \rightarrow \lim_{n\to\infty} d[n] = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

پس در این حالت، در ابتدا خسرو ۳ واحد به شیرین علاقه‌مند است و شیرین ۵ واحد به خسرو و به گذشت زمان محبت این دو شخص باهم یکسان می‌شود و هر دو ۴ واحد به یکدیگر علاقه‌مند خواهند بود. همچنین اختلاف علاقه‌شان به یکدیگر در هر گام نصف خواهد شد.

ح)

$$Av = \lambda v \rightarrow (A - \lambda I)v = 0 \rightarrow |A - \lambda I| = 0 \rightarrow \begin{vmatrix} 1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} = 0 \rightarrow (1-\lambda)^2 - 1 = 0 \rightarrow \lambda_1 = 0, \lambda_2 = 2$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0 \rightarrow \begin{bmatrix} x+y \\ x+y \end{bmatrix} = 0 \rightarrow x = -y \rightarrow v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0 \rightarrow \begin{bmatrix} -x+y \\ x-y \end{bmatrix} = 0 \rightarrow x = y \rightarrow v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

ط) این حالت هم مشابه قسمت‌های قبل انتظار می‌رود با ضرب‌های A در بردار ویژه به همان بردار ویژه با مقدار ویژه متناظر با آن برسیم:

$$d[1] = Ad[0] = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ -x \end{bmatrix} = \begin{bmatrix} x-x \\ x-x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$d[i] = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \lim_{n\to\infty} d[n] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

ی)

$$d[1] = Ad[0] = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2d[0]$$

$$d[i] = 2^i d[0] \rightarrow \lim_{n\to\infty} d[n] = \lim_{n\to\infty} 2^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \lim_{t\to\infty} \begin{bmatrix} t \\ t \end{bmatrix}$$

لذا در این حالت دوستی خسرو و شیرین هربار دو برابر می‌شود و به بی‌نهایت میل پیدا می‌کند!

ک)

$$Av = \lambda v \rightarrow (A - \lambda I)v = 0 \rightarrow |A - \lambda I| = 0 \rightarrow \begin{vmatrix} 1-\lambda & -2 \\ -2 & 1-\lambda \end{vmatrix} = 0 \rightarrow (1-\lambda)^2 - (-2)^2 = 0 \rightarrow \lambda_1 = -1 , \lambda_2 = 3$$

$$\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0 \rightarrow \begin{bmatrix} 2x-2y \\ -2x+2y \end{bmatrix} = 0 \rightarrow x = y \rightarrow v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} -2 & -2 \\ -2 & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 0 \rightarrow \begin{bmatrix} -2x-2y \\ -2x-2y \end{bmatrix} = 0 \rightarrow x = -y \rightarrow v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

ل)

$$d[1] = Ad[0] = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x \\ -x \end{bmatrix} = \begin{bmatrix} x + 2x \\ -2x - x \end{bmatrix} = \begin{bmatrix} 3x \\ -3x \end{bmatrix} = 3 \begin{bmatrix} x \\ -x \end{bmatrix}$$

$$d[i] = 3^i d[0] \to \lim_{n\to\infty} d[n] = \lim_{n\to\infty} 3^n \begin{bmatrix} x \\ -x \end{bmatrix} = \lim_{t\to\infty} \begin{bmatrix} t \\ -t \end{bmatrix}$$

در این مورد اگر در ابتدا خسرو شیرین را دوست داشته باشد و شیرین به همان میزان از خسـرو بدش بیاید، در هر بار هم علاقه و هم نفرت سـه برابر می‌شـود تا در نهایت خسـرو بی‌نهایت شـیرین را دوسـت داشـته باشـد و شـیرین بی‌نهایت از خسـرو متنفر شـود! به طور مشـابه اگر در ابتدا شـیرین خسـرو را مقداری دوسـت داشـته باشـد و به همان میزان خسـرو از شـیرین متنفر باشـد در نهایت شـیرین بی‌نهایت به خسـرو علاقه‌مند خواهد بود درحالی که خسـرو بی‌نهایت از شـیرین متنفر خواهد بود. اگر در ابتدا هر دو هیچ‌حسـی نسـبت به همدیگر نداشـته باشـند تا ابد و در هرلحظه این بی‌حسی ادامه‌دار خواهد بود.

م)

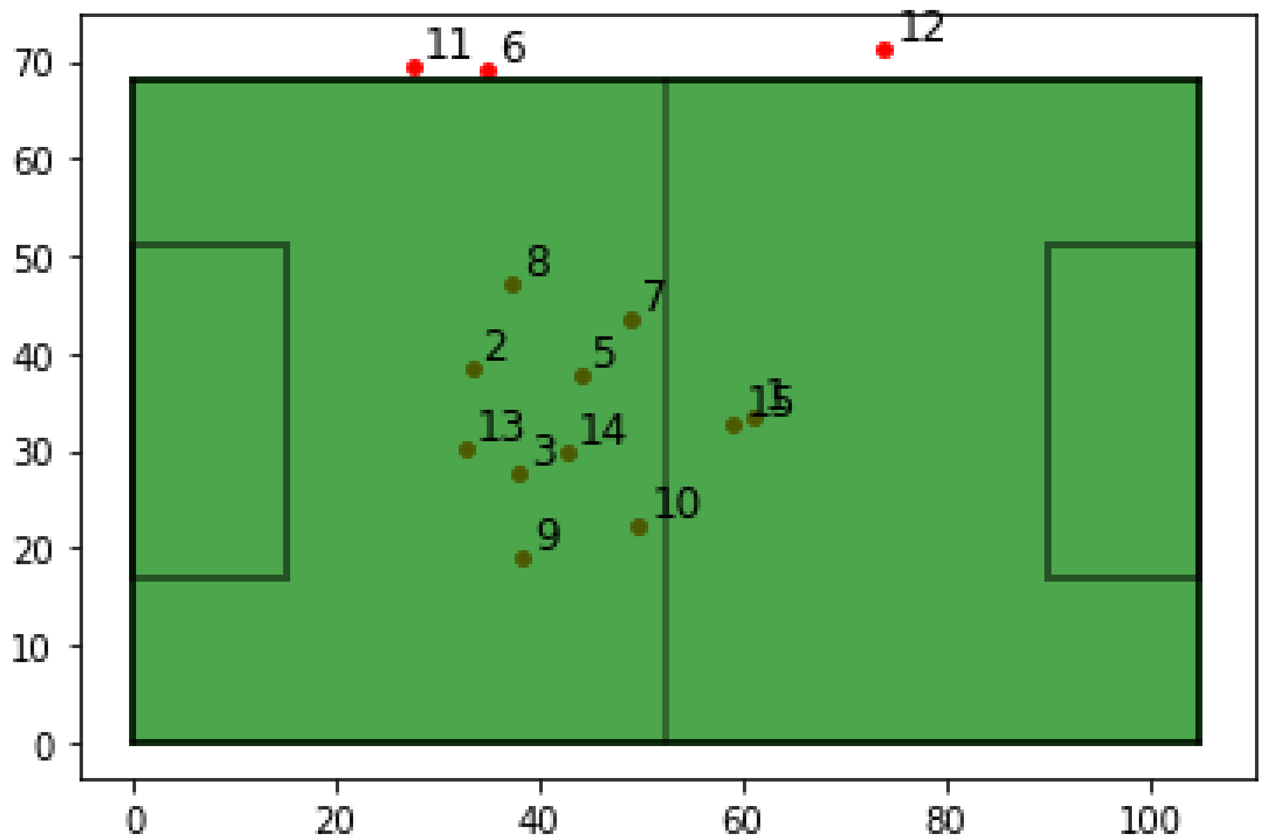
$$d[1] = Ad[0] = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} = -\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$d[i] = (-1)^i d[0] \to \lim_{n\to\infty} d[n] = \lim_{n\to\infty} (-1)^n \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

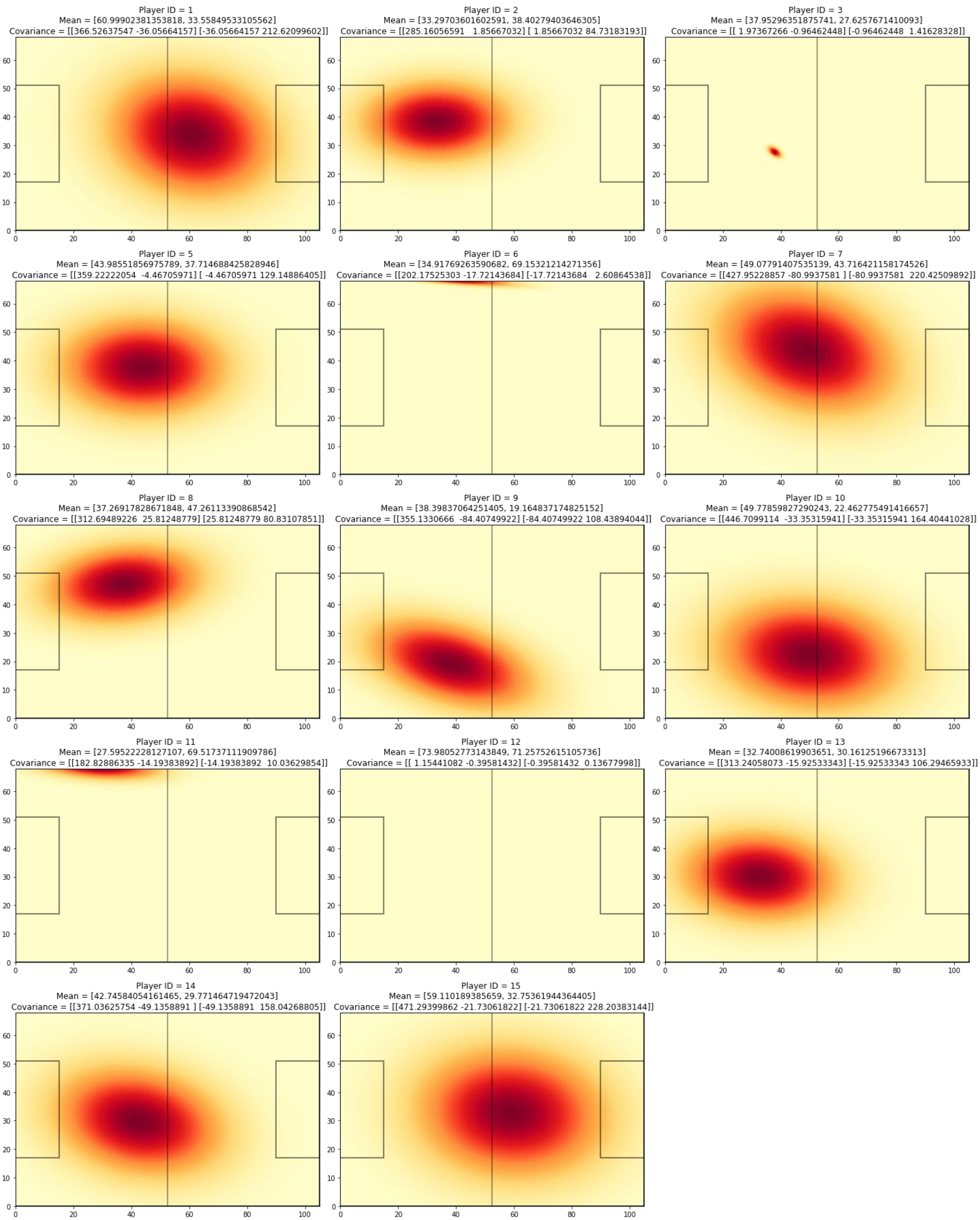
در این حالت در ابتدا خسرو و شیرین همدیگر را به یک میزان دوست دارند. در لحظه بعد به همان میزان از یکدگیر بدشـان می‌آید و مجددا باهم دوسـت می‌شـوند و تا انتها یک لحظه در میان باهم دوست می‌شوند و قهر می‌کنند!

سوال ۶

الف) در فایل موجود چهارده بازیکن دیده می‌شـــوند. به نظر ســـه بازیکن ۶، ۱۱ و ۱۲ بازیکنان ذخیره هستند ولی با این حال ترسیم شده اند.



ب)

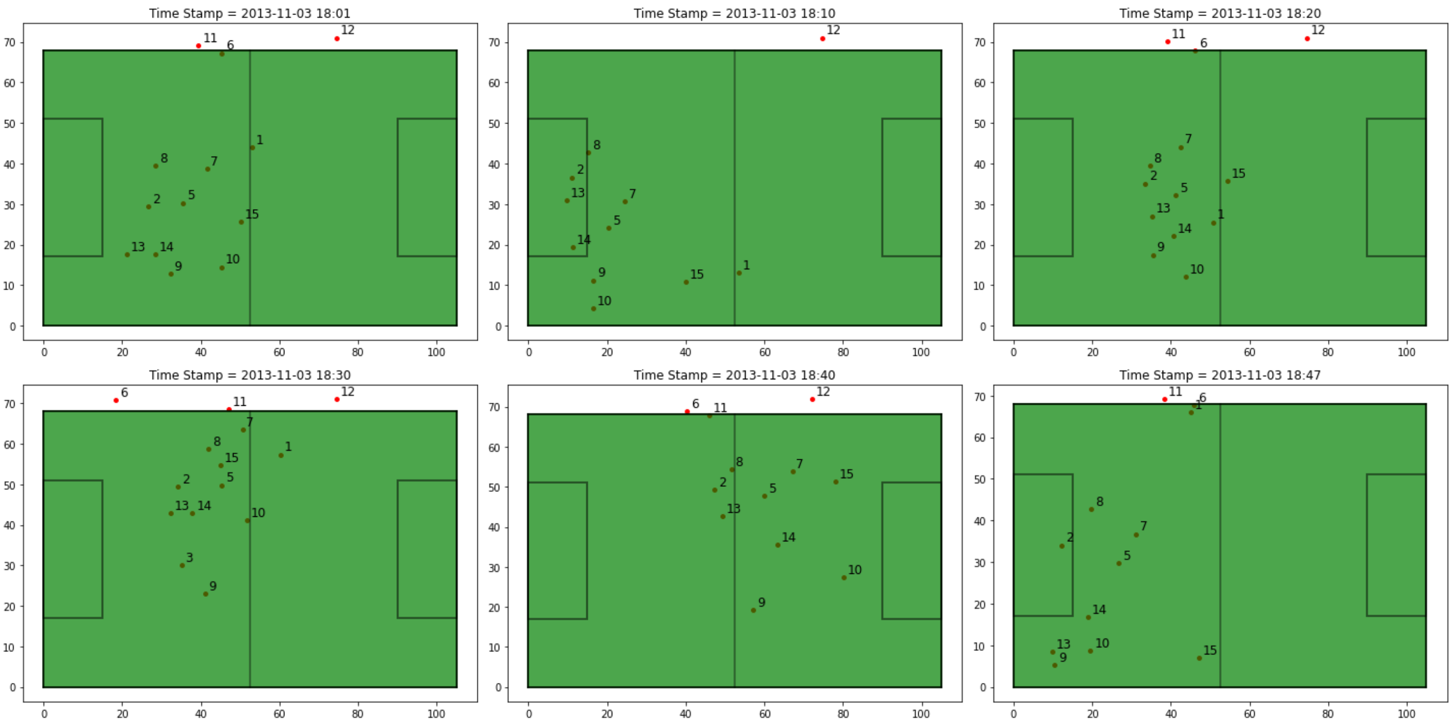

Player ID = 1
Mean = [60.99902381353818, 33.55849533105562]
Covariance = [[366.52637547 -36.05664157] [-36.05664157 212.62099602]]
Player ID = 2
Mean = [33.29703601602591, 38.40279403646305]
Covariance = [[285.16056591 1.85667032] [1.85667032 84.73183193]]
Player ID = 3
Mean = [37.95296351875741, 27.6257671410093]
Covariance = [[1.97367266 -0.96462448] [-0.96462448 1.41628328]]
Player ID = 5
Mean = [43.98551856975789, 37.714688425828946]
Covariance = [[359.22222054 -4.46705971] [-4.46705971 129.14886405]]
Player ID = 6
Mean = [34.91769263590682, 69.15321214271356]
Covariance = [[202.17525303 -17.72143684] [-17.72143684 2.60864538]]
Player ID = 7
Mean = [49.07791407535139, 43.716421158174526]
Covariance = [[427.95228857 -80.9937581] [-80.9937581 220.42509892]]
Player ID = 8
Mean = [37.26917828671848, 47.26113390868542]
Covariance = [[312.69489226 25.81248779] [25.81248779 80.83107851]]
Player ID = 9
Mean = [38.39837064251405, 19.164837174825152]
Covariance = [[355.1330666 -84.40749922] [-84.40749922 108.43894044]]
Player ID = 10
Mean = [49.77859827290243, 22.462775491416657]
Covariance = [[446.7099114 -33.35315941] [-33.35315941 164.40441028]]
Player ID = 11
Mean = [27.59522228127107, 69.51737111909786]
Covariance = [[182.82886335 -14.19383892] [-14.19383892 10.03629854]]
Player ID = 12
Mean = [73.98052773143849, 71.25752615105736]
Covariance = [[1.15441082 -0.39581432] [-0.39581432 0.13677998]]
Player ID = 13
Mean = [32.74008619903651, 30.16125196673313]
Covariance = [[313.24058073 -15.92533343] [-15.92533343 106.29465933]]
Player ID = 14
Mean = [42.74584054161465, 29.771464719472043]
Covariance = [[371.03625754 -49.1358891] [-49.1358891 158.04268805]]
Player ID = 15
Mean = [59.110189385659, 32.75361944364405]
Covariance = [[471.29399862 -21.73061822] [-21.73061822 228.20383144]]

ج)

	Player ID	Point	Probability
0	1	(92.32, 53.69)	4.157628e-05
1	1	(48.06, 7.46)	7.624918e-05
2	1	(89.89, 44.33)	1.181903e-04
3	15	(79.1, 26.65)	2.999878e-04
4	15	(39.49, 27.76)	2.993977e-04
5	15	(25.72, 26.24)	1.292443e-04
6	11	(3.15, 57.17)	9.007840e-09
7	11	(103.56, 64.18)	5.412237e-10
8	11	(30.91, 34.42)	1.189301e-32

د) برای تحلیل نقش بازیکنان، علاوه بر مراجعه به نمودارهای قسمت قبل، چندین تصویر در لحظات مختلف بازی را هم ترسیم کردیم. به علاوه به انرژی و فاصله پیموده‌شده نهایی توجه داشتیم.

بازیکن شماره ۶، ۱۱ و ۱۲ باتوجه به میانگین موقعیت و مجموع فاصله طی کرده شان بازیکنان ذخیره بودند. بازیکن شماره ۳ ظاهرا بازی نکرده است یا شاید هم برای آن داده زیادی جمع نشده است. هم اینکه تنها از دقیقه ۲۶ تا ۳۲ برایش داده وجود دارد و هم اینکه در کل این مدت در یک محل قرار داشته است. با کنار گذاشتن این بازیکنان، ده بازیکن باقی می‌مانند که باتوجه به نمودارها، احتمالا دروازه‌بان ترسیم نشده است؛ قسمت چپ زمین طرف خودی است و سیستم بازی به شکل ۲-۴-۴ است. در جدول زیر می‌توانید نقش‌های پیشنهادی من را مشاهده بفرمایید.



شماره بازیکن	نقش پیشنهادی
۱	مهاجم
۲	دفاع وسط
۳	؟!
۵	هافبک وسط
۶	ذخیره
۷	هافبک چپ
۸	دفاع چپ
۹	دفاع راست
۱۰	هافبک راست
۱۱	ذخیره
۱۲	ذخیره
۱۳	دفاع وسط
۱۴	هافبک وسط
۱۵	مهاجم

سوال ۷[4][5][6]

الف) Whitening Transform یک تبدیل خطی است که آرایه‌ای از متغیرهای تصادفی با یک ماتریس کواریانس را به آرایه‌ای از متغیرهای تصـــادفی تبدیل می‌کند که دیگر بین آن‌ها هیچ همبســتگی وجود ندارد و واریانس هر متغیر تصــادفی هم برابر با یک خواهد شـــد. پس با درنظر گرفتن این تبدیل به عنوان یک گام پیش‌پردازش می‌توان متغیرهای تصـــادفی را از یکدگیر جدا کرد. چنین چیزی باعث می‌شـــود در برخی از الگوریتم‌های یادگیری ماشـــین آموزش بهتری داشـــته باشـــیم. (به همان دلیل که ویژگی‌های هم‌بسته زائد است!)

ب) شـاید بتوان گفت فضـای پیوسـته حالت عام فضـای گسـسـته اسـت! رگرسـیون را می‌توان حالت کلی‌تری از دسته‌بندی دانست چون می‌توان برای هر کلاس یک نقطه در فضای پیوسته درنظر گرفت؛ لذا در این صورت یک الگوریتم رگرسیون می‌تواند هر داده را به نزدیک‌ترین یکی از نقطه‌های متناظر با کلاس‌ها در فضـــای پیوســـته و کلاس آن نســبت دهد. یا شــاید بتوان برای هر کلاس یک بعد پیوســته در نظر گرفت و احتمال تعلق به آن کلاس را با آن بعد نشـــان داد. طبیعتا هر داده به کلاســـی تعلق دارد که مقدار آن بعدش بیشتر باشد. و سایر ایده‌های مشابه...

ج) یک بردار ویژه برای یک ماتریس M، تبدیلی غیرصـــفر اســـت که اگر بر ماتریس M اعمال شود حاصلش نسبتی از همان ماتریس M شود. به آن نسبت هم، مقدار ویژه متناظر با آن بردار ویژه گفته می‌شود.

با تعریف رســـمی‌تر می‌توان گفت که اگر M یک تبدیل خطی از فضـــای برداری V به خودش در میدان F باشـد و v یک بردار غیرصـفر در فضـای v باشـد؛ آنگاه v یک بردار ویژه برای M است اگر $M(v)$ یک ضریب اسکالر از v باشد که می‌توان آن را با فرمول زیر نوشت. در این فرمول λ یک عدد اسکالر در میدان F است که مقدار ویژه متناسب با بردار ویژه v درنظر گرفته خواهد شد.

---

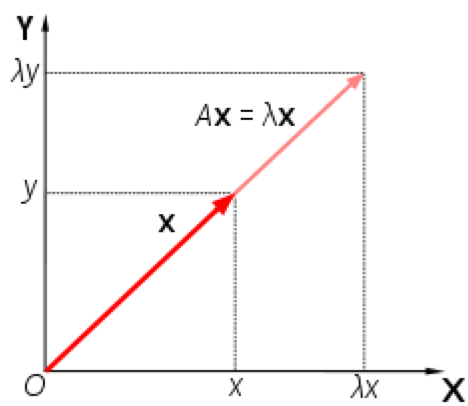
[4] https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors

[5] https://www.dhruvonmath.com/2019/03/20/pagerank/

[6] http://pi.math.cornell.edu/~mec/Winter2009/RalucaRemus/Lecture3/lecture3.html

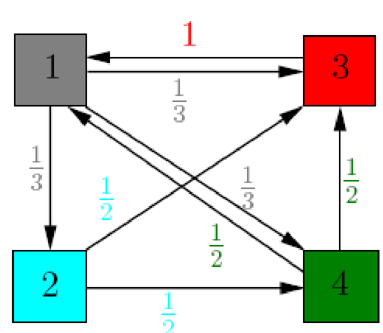
$$Mv = \lambda v$$

د) از نظر هندسی، یک بردار ویژه را می‌توان جهتی دانست که اگر یک تبدیل خطی بر آن اعمال شود و آن را بکشد، همچنان در همان جهت یا خلاف آن جهت باقی بماند و چرخش نداشته باشد. مقدار ویژه مقیاسی است که مشخص می‌کند با این تبدیل به چه میزان بردار ویژه کشیده شده است. همچنین منفی‌بودن مقدار ویژه نشان‌دهنده آن است که پس از اعمال تبدیل خطی، بردار ویژه در خلاف جهت اولیه قرار می‌گیرد. به عنوان مثال در تصویر زیر مشخص است که اگر x بردار ویژه تبدیل A باشد. پس از اعمال A بر آن، همچنان جهت x تغییر نکرده است. فقط مقدار آن عوض شده است که نسبت آن وابسته به مقدار ویژه یعنی λ است.



ه) گوگل برای آنکه صفحات را رتبه‌بندی کند از الگوریتم PageRank استفاده می‌کند. به طور خلاصه، متناسب با این الگوریتم اعتبار هر صفحه متناسب با اعتبار صفحاتی است که به آن ارجاع داشته‌اند. در این الگوریتم یک ماتریس N*N به نام A ساخته می‌شود که N تعداد صفحات است. در ماتریس A مشخص می‌شود که یک صفحه به یک صفحه دیگر ارجاع داده است یا خیر. نمونه‌ای از این ماتریس برای یک گراف آورده شده است:

$$\begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$



حال باید امتیازها را به گونه‌ای به ماتریس به گره‌ها تخصـیص داد که روابط برقرار بماند یعنی امتیاز هر راس برابر باشـد با امتیاز گره‌هایی که به آن ارجاع داده‌اند ضـرب در یک عدد نرمال‌سـازی. لذا اگر برداری از امتیاز رئوس را v بنامیم باید داشته باشیم:

$$Av = v$$

برای پیدا کردن v یک راه آن اسـت که ابتدا مقدار تصـادفی برای آن درنظر بگیریم و سـپس به تعداد بار زیاد A را بر آن اعمال کنیم تا به یک بردار همگرا شـــود و یک راه دیگر آن اســـت که بردار ویژه متناسب با مقدار ویژه ۱ برای ماتریس A را پیدا کنیم که روال سریع‌تری محسوب می‌شود.