**به نام خدا**

دانشگاه صنعتی امیرکبیر
( پلی تکنیک تهران )

# تمرین درس شناسایی آماری الگو

**فردین آیار**

**شماره دانشجویی: ۹۹۱۳۱۰۴۰**

**استاد: دکتر محمد رحمتی**

**دانشکده کامپیوتر- پاییز ۹۹**

# 1)

a1) regression

a2) هر سیستمی که برای ورود اطلاعات دانشجویان استفاده شود( پایگاه های داده، دستگاه اسکن پرسشنامه و ...)

a3) نمرات دانشجویان ترم های گذشته و سایر اطلاعات در صورت وجود

a4) داده های میتوانند از روش های مختلفی مانند پرسیدن از دانشجویان، دیتابیس های دانشگاه(در صورت امکان) و ... جمع آوری شوند.

a5) تعداد واحدهای دانشجویان در ترم، معدل دانشجویان، میزان مطالعه(در صورت وجود اطلاعات)، نمرات درس های مشابه و ...

a6) برخی داده ها ممکن است موجود نباشد( به عنوان میزان مطالعه برخی از دانشجویان) در این صورت باید با استفاده از روش هایی مانند مقدار ثابت،استفاده از میانگین و ... داده های ناموجود را جاگذاری یا در صورت امکان نمونه را حذف کرد.

بعضی داده ها ممکن است پرت باشند، به عنوان مثال دانشجویی که با معدل کم، نمره بسیار بالایی کسب کرده است. در این صورت ممکن است داده ها نیاز به حذف یا تغییر داشته باشند.

a7) بعضی ویژگی های بسیار مهم ممکن است قابل اعمال نباشند. به عنوان مثال معیاری برای سنجش استعداد دانشجویان به صورت مستقیم وجود ندارد. حال آنکه این ویژگی در نمره دانشجویان تاثیر بسیاری دارد. همچنین بعضی سوابق ممکن است در نظر گرفته نشود؛ به عنوان مثال ممکن است دانشجویان قبلا درس مشابهی گذرانده باشند و یا در مورد موضوع درس، مطالعه قبلی داشته باشند.

a8) توانایی درنظر گرفتن چندین معیار و بررسی نمونه های آموزشی بسیار زیاد از مهم ترین نقاط قوت این سیستم است. از طرفی در مقایسه با انسان، چنین سیستمی توانایی در نظر گرفتن استثناها و معیارهای غیرکمی(مانند استعداد) را ندارد.

b1) regression

b2) هر سیستمی که برای ورود اطلاعات استفاده شود.

b3) داده های سال های گذشته و پارامترهای اقتصادی مربوط به آن سالها

b4) استفاده از داده های موجود در اینترنت، جمع آوری داده ها از مراجع رسمی

b5) نرخ تورم، نرخ بهره،رشد نقدینگی، میزان صادرات/واردات کشور، سایر پارامترهای اقتصادی(بدهی دولت، GDP و ...)

b6) همانطور که در بخش a6 به صورت کامل توضیح داده شد، بعضی داده ها ممکن است وجود نداشته باشند یا پرت باشند. در این صورت تغییرات لازم در دیتاست باید اعمال شود.

علاوه بر این نرمال سازی داده ها ممکن است مورد نیاز باشد، به عنوان مثال دو پارامتر ((نرخ تورم)) و ((میزان صادارت کشور)) بازه های بسیار متفاوتی دارند که ممکن است تاثیر منفی در الگوریتم داشته باشد.

b7) مهمترین عاملی که باعث خطا در سیستم می شود، تاثیر معیارهایی هستند که قابل کمی سازی نیستند. به عنوان مثال عوامل سیاسی اگرچه تاثیر خود را روی پارامترهایی مانند ((میزان صادرات)) کشور نشان می دهند. اما خود به عنوان یک عامل موثر باید مستقلا در نظر گرفته شود؛ در صورتی که چندان در قالب اعداد توصیف پذیر نیست. سایر پارامترهای موثر مانند ((انتظارات تورمی)) نیز قابلیت اندازه گیری ندارند.

از طرف دیگر شرایط پیش بینی نشده مانند جنگ، فجایای طبیعی و... می توانند خروجی سیستم را غیر معتبر کنند.

b8) مانند سیستم قبلی، مهمترین مزیت این سیستم استفاده از معیارها و داده های بسیار زیاد است. اما رابطه بسیار زیاد مسائل اقتصادی و سیاسی، مسائل پیش بینی نشده و همچنین تاثیر رفتار جمعی جامعه، مواردی هستند که باعث ضعف سیستم ما در برابر پیشبینی های انسانی می شوند.

c1) classification اگر خصوصیت های شخصیتی مدنظر تعریف شده و مشخص باشند. Clustering اگر دسته بندی کلی مدنظر باشد یا انواع شخصیت ها تعریف مشخصی نداشته باشند.

c2) هر سیستمی که برای ورود اطلاعات استفاده شود. ( اسکنر پرسشنامه و ...)

c3) دیتاست های موجود در وب یا آزمایش های روانشناختی قبلی

c4) با استفاده از نظر سنجی از دانشجویان،تحلیل داده های موجود در صفحات اجتماعی دانشجویان

c5) میزان ارتباطات اجتماعی، کلمات کلیدی به کار رفته در شبکه های اجتماعی، سبک زندگی و عادات کلی روزانه( میزان مطالعه و ...)، تعداد دوستان، نمرات درسی و سایر داده هایی که از نظرسنجی میتوان استخراج کرد ( ویژگی های کلی شخصیتی ، باور ها و عقاید و...)

c6) داده های شبکه های اجتماعی نیاز به پیش پردازش برای استخراج کلمات کلیدی و ... دارند. سایر پیش پردازش ها مانند سیستم های قبلی است.

c7) داده های روانشناختی معمولا توصیفی هستند وقابلیت اندازه گیری پایینی دارند. از طرفی برخی شخصیت های درونگرا را نمی توان بر اساس داده های شبکه اجتماعی یا نظرسنجی تحلیل کرد و به طور کلی داده های زیادی برای دسته بندی شخصیتی مورد نیاز است که ممکن است در دسترس نباشند.

c8) همانطور که در بخش قبلی توضیح داده شد، داده های روانشناختی و شخصیتی قابلیت اندازه گیری بسیار کمی دارند. همچنین ماشین توانایی درک احساسات انسانی را ندارد. از طرفی توانایی پردازش داده های شبکه های اجتماعی که حجم زیادی دارند، برتری چنین سیستمی نسبت به انسان است.

## 2)

a) رنگ پوست- سایز بینی – سایز لب ها – شکل و محل قرارگیری چشم ها

b) میزان انحنای لب ها – چین و چروک اطراف دهان و چشم – مشخص بودن یا نبودن دندان ها

c) نسبت ابعاد اجزای صورت به مساحت صورت- پرپشتی ابرو ها- لکه ها و چین و چروک صورت- شکل و فرم گونه ها –شفافیت پوست-ابعاد صورت

d) شفافیت پوست- موهای صورت- عرض چهره-شکل چانه و انحنای صورت-ضخامت ابروها

e) شکل کلی چهره – رنگ چهره – اجزای چهره ، نسبت ، سایز و فاصله آنها

# ۳)

a) با توجه به تعریف تابع چگالی احتمال داریم:

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \xRightarrow{0\leq x\leq 1} \int_0^1 (cx)dx = 1 \Rightarrow \frac{cx^2}{2}|_0^1 = 1 \Rightarrow \frac{c}{2} = 1 \Rightarrow \boldsymbol{c = 2}$$

b)

$$P(0 \leq X \leq 0.5) = \int_0^{0.5} f(x)dx \Rightarrow \int_0^1 (2x)dx = x^2|_0^{0.5} = 0.25$$

c)

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^1 (2x^2)dx = \frac{2x^3}{3}|_0^1 = \frac{2}{3}$$

d)

$$\begin{aligned} Var(X) &= E((X-\mu)^2) \\ &= \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = \int_0^1 \left(x-\frac{2}{3}\right)^2 (2x)dx = \int_0^1 (2x^3 - \frac{8}{3}x^2 + \frac{8}{9}x)dx \\ &= \frac{1}{2}x^4 - \frac{8}{9}x^3 + \frac{4}{9}x^2|_0^1 = \frac{1}{2} - \frac{8}{9} + \frac{4}{9} = \frac{1}{18} \end{aligned}$$

e) روش اول:

$$E(2X-2) = \int_{-\infty}^{+\infty} (2x-2)f(x)dx = \int_0^1 (4x^2-4x)dx = \frac{4x^3}{3} - 2x^2|_0^1 = -\frac{2}{3}$$

روش دوم:

$$E(2X-2) = 2E(X) - E(2) = 2 \times \frac{2}{3} - 2 = -\frac{2}{3}$$

f) روش اول:

$$\begin{aligned} Var(2X-2) &= E(((2X-2)-\mu_{2X-2})^2) \\ &= \int_{-\infty}^{+\infty} ((2x-2)-\mu_{2X-2})^2 f(x)dx \\ &= \int_0^1 \left(2x-2+\frac{2}{3}\right)^2 (2x)dx = \int_0^1 (8x^3 - \frac{32}{3}x^2 + \frac{32}{9}x)dx \\ &= 2x^4 - \frac{32}{9}x^3 + \frac{16}{9}x^2|_0^1 = 2 - \frac{32}{9} + \frac{16}{9} = \frac{2}{9} \end{aligned}$$

روش دوم:

$$Var(2X-2)=4Var(X)=4\times\frac{1}{18}=\frac{2}{9}$$

g) برای استفاده جدول توزیع نرمال تعریف می کنیم $z=\frac{X-\mu}{\sigma}=\frac{X-1}{3}$

$$P(-2\le X\le 1)=P\left(\frac{-2-1}{3}\le z\le\frac{1-1}{3}\right)=P(-1\le z\le 0)=P(z\le 0)-P(z\le -1)$$
$$=0.5-0.1587=0.3413$$

h) اگرچه پارامترهای خواسته شده همان $\sigma$ و $\mu$ هستند اما مجدداً آن ها را به کمک تعاریف بدست می آوریم. در اینجا مسئله به صورت پارامتری حل شده و مقادیر $\sigma$ و $\mu$ را جهت سادگی جاگذاری نمی کنیم.

$$E(X)=\int_{-\infty}^{+\infty}xf(x)dx=\int_{-\infty}^{+\infty}\frac{x}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-1}{\sigma}\right)^2}dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}x\times e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$$

در نظر میگیریم:

$$\begin{cases} y=\dfrac{x-\mu}{\sqrt{2}\sigma} \\ x=\sqrt{2}\sigma y+\mu \\ dy=\dfrac{dx}{\sqrt{2}\sigma} \\ dx=\sqrt{2}\sigma\times dy \end{cases}$$

بنابراین:

$$E(X)=\frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(\sqrt{2}\sigma y+\mu\right)\times e^{-y^2}dy \qquad (I)$$

انتگرال فوق را * می نامیم:

$$*=\sqrt{2}\sigma\int_{-\infty}^{+\infty}y\times e^{-y^2}dy+\mu\int_{-\infty}^{+\infty}e^{-y^2}dy$$

در عبارت بالا، عبارت درون انتگرال اول یک تابع فرد است، بنابراین طبق تعریف تابع فرد، حاصل انتگرال صفر می شود. همچنین، انتگرال دوم ، انتگرال گاووسی نام دارد. محاسبه این انتگرال معروف، با روش های عادی دشوار است، اما می توان نشان داد که حاصل آن $\sqrt{\pi}$ است. بنابراین $*=\sqrt{\pi}\mu$ و با جاگذاری در $(I)$ داریم:

$$E(X)=\frac{1}{\sqrt{\pi}}\times\sqrt{\pi}\mu=\mu$$

برای محاسبه واریانس داریم:

$$Var(X)=E((X-\mu)^2)=\int_{-\infty}^{+\infty}(x-\mu)^2f(x)dx=\int_{-\infty}^{+\infty}\frac{(x-\mu)^2}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$$

حال تغییر متغییر زیر را اعمال می کنیم:

$$\begin{cases} y = x - \mu \\ dy = dx \\ x = \mu + y \end{cases}$$

$$Var(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 \times e^{-\frac{y^2}{2\sigma^2}} dy$$

برای حل انتگرال فوق با استفاده از روش جز به جز داریم:

$$\begin{cases} u = y \\ du = dy \\ dv = y^2 \times e^{-\frac{y^2}{2\sigma^2}} dy \\ v = -\sigma^2 e^{-\frac{y^2}{2\sigma^2}} \end{cases}$$

$$Var(X) = \frac{1}{\sigma\sqrt{2\pi}} \left[ -\sigma^2 y e^{-\frac{y^2}{2\sigma^2}} |_{-\infty}^{+\infty} + \sigma^2 \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2\sigma^2}} dy \right] = 0 + \sigma^2 \int_{-\infty}^{+\infty} \frac{e}{\sigma\sqrt{2\pi}}^{-\frac{y^2}{2\sigma^2}} dy$$

تابع درون انتگرال، توزیع نرمال با واریس $\sigma$ و میانگین صفر است. می دانیم حاصل انتگرال توابع توزیع احتمال در بازه منفی بینهایت تا مثبت بی نهایت ۱ است. بنابراین:

$$Var(X) = \sigma^2$$

*i*) ابتدا تابع توزیع تجمعی $Y$ را می یابیم. سپس با مشتق گرفتن از آن، به تابع چگالی احتمال $Y$ می رسیم:

$$F_Y(t) = P(Y \le t) = P(2X - 1 \le t) = P\left(X \le \frac{t+1}{2}\right) = \int_{-\infty}^{\frac{t+1}{2}} \frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-1}{3}\right)^2} dx$$

برای محاسبه مشتقِ انتگرال بالا با استفاده از قضیه لایب نیتز داریم:

$$f_y(t) = \frac{d}{dt} F_y(t) = \frac{1}{2} \frac{1}{\sqrt{2\pi} \times 3} e^{-\frac{1}{2}(\frac{\frac{t+1}{2}-1}{3})^2} = \frac{1}{\sqrt{2\pi} \times 6} e^{-\frac{1}{2}(\frac{\frac{t+1}{2}-1}{3})^2} = \frac{1}{\sqrt{2\pi} \times 6} e^{-\frac{1}{2}(\frac{t-1}{6})^2}$$

بنابراین E(Y)=1 و Var(Y)=36 می باشد.

برای اثبات درستی اعداد به دست آمده با توجه به خواص واریانس و امید ریاضی داریم:

$$\begin{cases} E(2X - 1) = 2E(X) - E(1) = 2 \times 1 - 1 = 1 \\ Var(2X - 1) = 4Var(X) = 4 \times 9 = 36 \end{cases}$$

j) احتمال خواسته شده، مطابق با توزیع دو جمله ای است که طبق آن داریم:

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

در عبارت بالا $n$ تعداد آزمایش، $k$ تعداد موفقیت و $p$ احتمال موفقیت هر آزمایش است. با توجه به صورت مسئله:

$$p(4) = \binom{7}{4} \left(\frac{1}{5}\right)^4 \times \left(\frac{4}{5}\right)^3 = 0.0286$$

k) با توجه به تعریف میانه ($m$) در متغیر های پیوسته با تابع توزیع احتمال f(x) داریم:

$$\int_{-\infty}^{m} f(x)dx = \frac{1}{2} \Rightarrow \int_{0}^{m} \frac{1}{4}(4 - x^2) = x - \frac{x^3}{12}\Big|_0^m = m - \frac{m^3}{12} = \frac{1}{2} \Rightarrow m - \frac{m^3}{12} - \frac{1}{12} = 0$$

با حل معادله بالا تنها جواب قابل قبول ( m در بازه ۰ تا ۲) ، *0.08338* است.

l)  مطابق سوال *g* عمل می کنیم:

$$P(X \leq 9.8) = P\left(z \leq \frac{9.8 - 10}{0.1}\right) = P(z \leq -2) = 0.0228$$

m)

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_{0}^{1} \left(\frac{4x}{\pi(1 + x^2)}\right) dx$$

برای حل انتگرال فوق از روش تغییر متغیر استفاده می کنیم.

$$\begin{cases} y = 1 + x^2 \\ dy = 2xdx \end{cases}$$

$$E(x) = \int_{1}^{2} \frac{2}{\pi y} dy = \frac{2\ln y}{\pi}\Big|_1^2 = \frac{2\ln 2}{\pi} = 0.441$$

## ۴)

a1)

$$A \times v = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \times \begin{bmatrix} -3 \\ -3 \end{bmatrix} = \begin{bmatrix} -15 \\ -15 \end{bmatrix} = 5 \begin{bmatrix} -3 \\ -3 \end{bmatrix} \quad v\ is\ eigenvector\ of\ A\ and\ 5\ is\ eigenvalue$$

a2)

$$A \times v = \begin{bmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{bmatrix} \times \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 19 \\ -10 \\ -1 \end{bmatrix} \quad v\ is\ not\ eigenvector\ of\ A$$

b1)

$$A \times v = \gamma \times v \Rightarrow (A - \gamma I)v = 0$$

$$\begin{bmatrix} 1-(-2) & -6 \\ -3 & 4-(-2) \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix} = 0 \xRightarrow{\text{elementary row operations}} \begin{bmatrix} 3 & -6 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow 3x - 6y = 0 \Rightarrow \begin{cases} x = 2t \\ y = t \end{cases}$$

$$v = t \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

بنابراین بردار $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ یک پایه برای فضای ویژه ماتریس Á است.

b2)

$$\begin{bmatrix} -2-6 & 4 & 2 \\ 2 & 1-6 & -2 \\ 4 & -2 & 5-6 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \xRightarrow{\text{ERO}} \begin{bmatrix} -8 & 4 & 2 \\ 2 & -5 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \Rightarrow \begin{cases} -8x + 4y + 2z = 0 \\ 2x - 5x - 2z = 0 \end{cases} \Rightarrow \begin{cases} x = t \\ y = -6t \\ z = 16t \end{cases}$$

$$\Rightarrow \qquad v = t \begin{bmatrix} 1 \\ -6 \\ 16 \end{bmatrix}$$

c1) ابتدا به کمک معادله زیر مقادیر ویژه و سپس بردار های ویژه را می یابیم:

$$A \times v = \gamma \times v \Rightarrow (A - \gamma I)v = 0 \Rightarrow |A - \gamma I| = 0$$

$$\begin{vmatrix} 3-\gamma & 2 \\ -1 & 6-\gamma \end{vmatrix} = 0 \Rightarrow (3-\gamma)(6-\gamma) + 2 = 0 \xRightarrow{solve\ for\ \gamma} \gamma = 5{,}4$$

$$\gamma = 5: \begin{bmatrix} -2 & 2 \\ -1 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix} = 0 \xRightarrow{ERO} \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow \begin{cases} x = t \\ y = -t \end{cases} \Rightarrow v = t \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\gamma = 4:\ \begin{bmatrix}-1 & 2\\ -1 & 2\end{bmatrix} \times \begin{bmatrix}x\\ y\end{bmatrix} = 0 \xRightarrow{ERO} \begin{bmatrix}-1 & 2\\ 0 & 0\end{bmatrix} \times \begin{bmatrix}x\\ y\end{bmatrix} = 0 \Rightarrow \begin{cases}x = 2t\\ y = t\end{cases} \Rightarrow v = t\begin{bmatrix}2\\ 1\end{bmatrix}$$

c2) با توجه به معادله $A \times v = \gamma \times v$ بدیهی است که مقدار ویژه صفر و بردار ویژه هر بردار دو بعدی می تواند باشد.

c3) مانند قسمت ۱ عمل می کنیم:

$$\begin{vmatrix}3-\gamma & 1 & -2\\ 2 & 3-\gamma & -2\\ 2 & 1 & -1-\gamma\end{vmatrix} = 0 \xRightarrow{solve\ for\ \gamma} \gamma = 1,3$$

$$\gamma = 1:\ \begin{bmatrix}2 & 1 & -2\\ 2 & 2 & -2\\ 2 & 1 & -2\end{bmatrix} \times \begin{bmatrix}x\\ y\\ z\end{bmatrix} = 0 \xRightarrow{ERO} \begin{bmatrix}2 & 1 & -2\\ 2 & 2 & -2\\ 0 & 0 & 0\end{bmatrix} \times \begin{bmatrix}x\\ y\\ z\end{bmatrix} = 0 \Rightarrow \begin{cases}x = t\\ y = 0\\ z = t\end{cases} \Rightarrow v = t\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}$$

$$\gamma = 3:\ \begin{bmatrix}0 & 1 & -2\\ 2 & 0 & -2\\ 2 & 1 & -4\end{bmatrix} \times \begin{bmatrix}x\\ y\\ z\end{bmatrix} = 0 \xRightarrow{ERO} \begin{bmatrix}0 & 1 & -2\\ 2 & 0 & -2\\ 0 & 0 & 0\end{bmatrix} \times \begin{bmatrix}x\\ y\\ z\end{bmatrix} = 0 \Rightarrow \begin{cases}x = t\\ y = 2t\\ z = t\end{cases} \Rightarrow v = t\begin{bmatrix}1\\ 2\\ 1\end{bmatrix}$$

d) با توجه به تعریف داریم:

$$\begin{cases}\begin{bmatrix}a & b\\ c & d\end{bmatrix} \times \begin{bmatrix}-\frac{3}{2}\\ 1\end{bmatrix} = \begin{bmatrix}\frac{3}{2}\\ -1\end{bmatrix}\\ \begin{bmatrix}a & b\\ c & d\end{bmatrix} \times \begin{bmatrix}\frac{1}{2}\\ 1\end{bmatrix} = \begin{bmatrix}\frac{7}{2}\\ 7\end{bmatrix}\end{cases} \Rightarrow \begin{cases}-\frac{3}{2}a + b = \frac{3}{2}\\ -\frac{3}{2}c + d = -1\\ \frac{1}{2}a + b = \frac{7}{2}\\ \frac{1}{2}c + d = 7\end{cases} \xRightarrow{solve\ (eq1,eq3)\&(eq2,eq4)} A = \begin{bmatrix}a = 1 & b = 3\\ c = 4 & d = 5\end{bmatrix}$$

e)

$$Av = \gamma v \xRightarrow{\times A^{-1}} A^{-1}Av = \gamma A^{-1}v \Rightarrow v = \gamma A^{-1}v \Rightarrow A^{-1}v = \gamma^{-1}v$$

f) ماتریسی $nxn$ diagonizable است اگر مجموع ((ابعاد)) eigenspace های آن برابر با n باشد. طبق این نکته ماتریسA ، diagonizable نیست.

g)

$$Av = \gamma v \xRightarrow{\times A} A^2v = \gamma Av \Rightarrow A^2v = \gamma^2 v$$

h) تعریف می کنیم :

$$\begin{cases}z_{n-1} = \begin{bmatrix}x_{n-1}\\ y_{n-1}\end{bmatrix}\\ M = \begin{bmatrix}2 & -3\\ -4 & 1\end{bmatrix}\end{cases} \Rightarrow z_n = M \times z_{n-1} \qquad (I)$$

حال به کمک بردار های ویژه ماتریس M را قطری می کنیم. در اینجا از ذکر محاسبات صرف نظر می کنیم. داریم:

$$\begin{cases} EigenVectors(M) = V = \begin{bmatrix} 0.7071 & 0.6 \\ -0.7071 & 0.8 \end{bmatrix} \\ EigenValues(M) = D = \begin{bmatrix} 5 & 0 \\ 0 & -2 \end{bmatrix} \end{cases}$$

در رابطه (I) ماتریس $M$ را با کمک بردارهای ویژه، به صورت $M = VDV^{-1}$ (این رابطه طبق تعریف مقدارهای ویژه به سادگی قابل اثبات است) قرار می دهیم:

$$z_n = M \times z_{n-1} = VDV^{-1}z_{n-1} = VDV^{-1}VDV^{-1}z_{n-2} = VD^2V^{-1}z_{n-2} = VD^nV^{-1}z_0$$

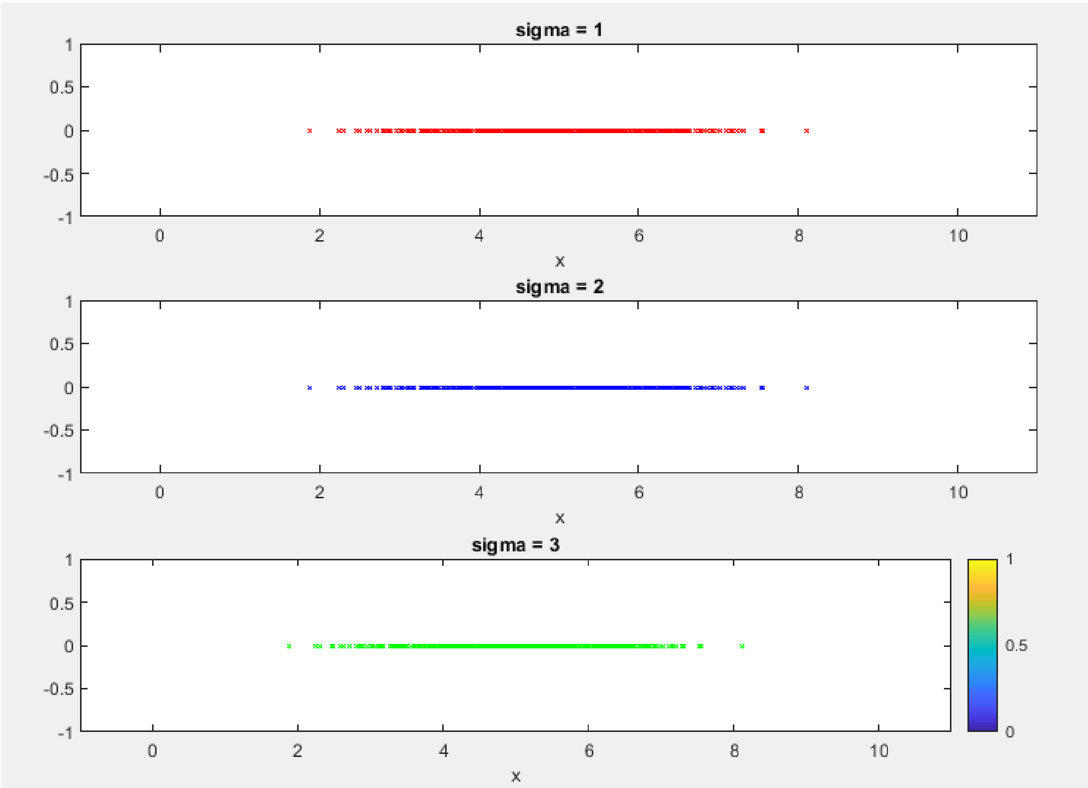
از آنجایی که ماتریس$D$ قطری است، توان $n$ ام آن به سادگی محاسبه می شود.

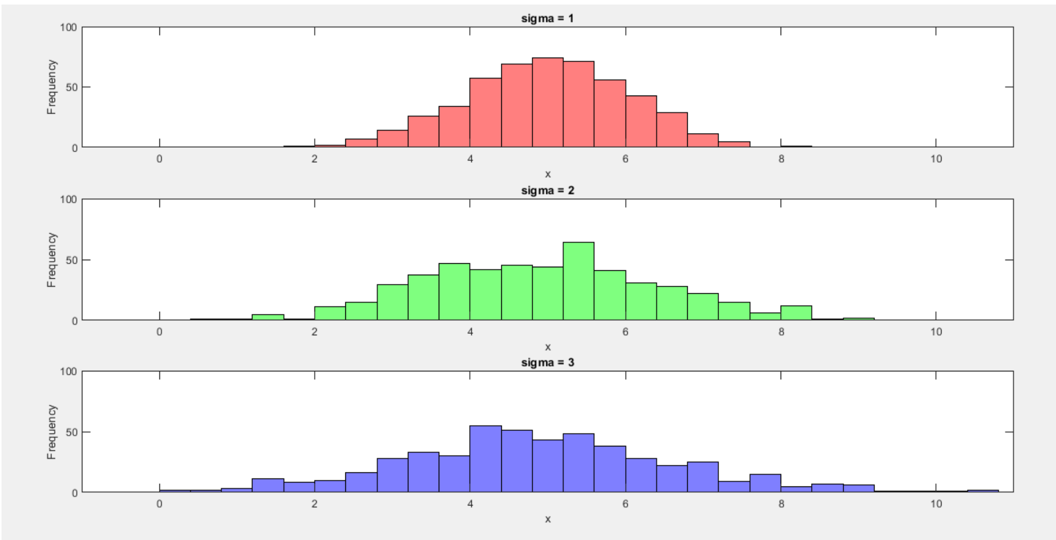
$$\begin{bmatrix} x \\ y \end{bmatrix} = VD^nV^{-1}z_0 = \begin{bmatrix} 0.7071 & 0.6 \\ -0.7071 & 0.8 \end{bmatrix} \times \begin{bmatrix} 5 & 0 \\ 0 & -2 \end{bmatrix}^n \times \begin{bmatrix} 0.8081 & -0.6061 \\ 0.7143 & 0.7143 \end{bmatrix} \times \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

## ۵)

a) اسکریپت a.m از برنامه متلب

در شکلهای زیر توزیع نقاط روی محور x برای سه حالت مختلف واریانس و همچنین هیستوگرام آن ها نشان داده شده است.

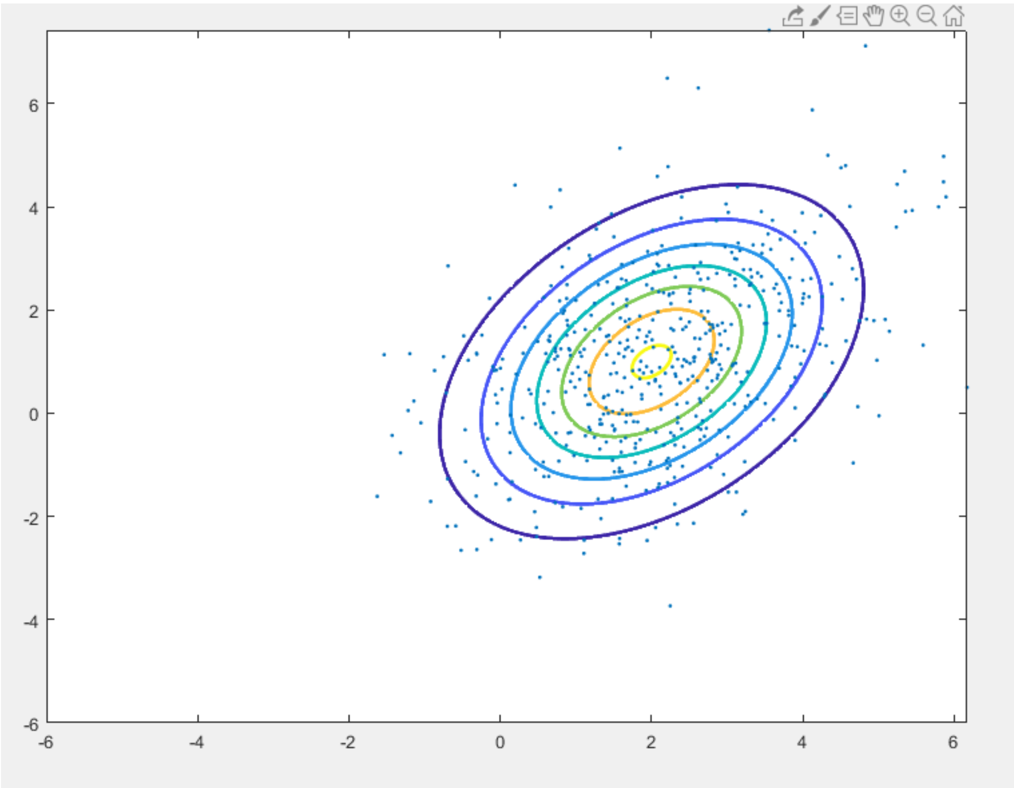


همانطور که در شکل های فوق مشاهده می‌شود، با افزایش سیگما، پراکندگی داده ها حول میانگین بیشتر شده است و فشردگی داده ها در بازه میانگین(۳)  کمتر شده است.

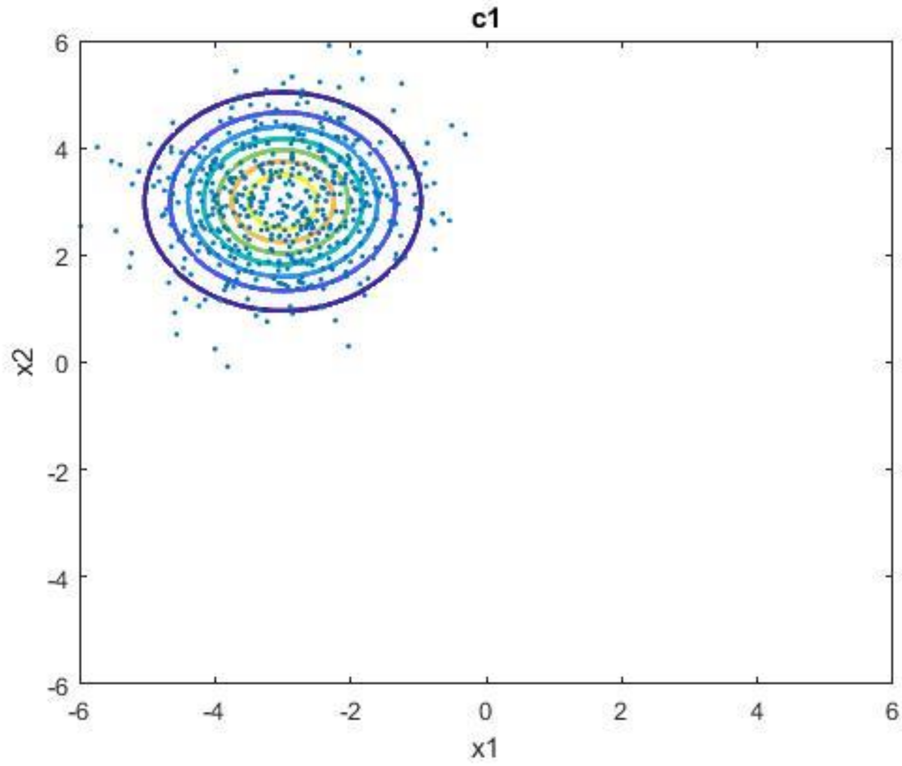
b) اسکریپت b.m از برنامه متلب

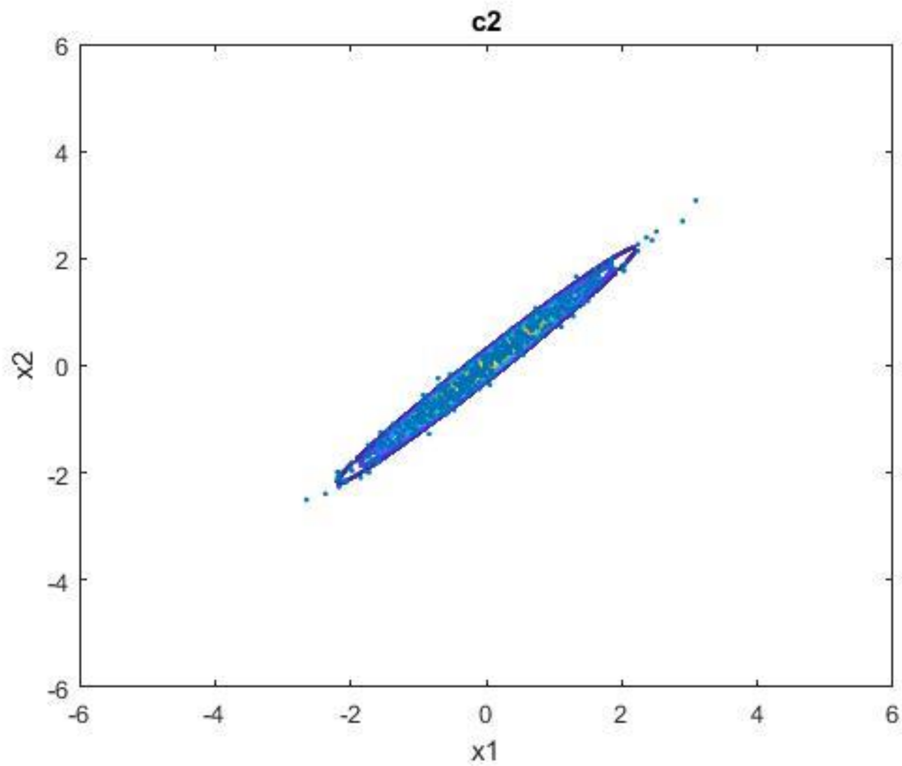
در شکل زیر داده های تولید شده و کانتورهای مربوطه رسم شده اند.

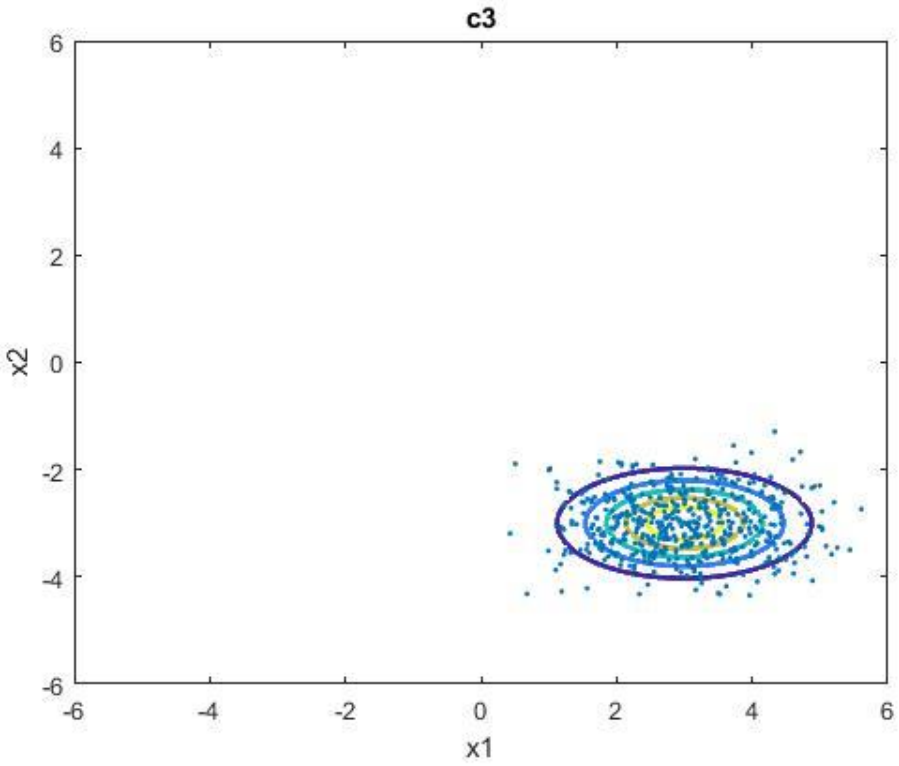



c)اسکریپت c.m از برنامه متلب

توزیع های خواسته شده در شکلهای زیر ارائه می شوند.


c1
x1
x2
-6
-4
-2
0
2
4
6


c2
x1
x2
-6
-4
-2
0
2
4
6




(d

$$M = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$\begin{cases} \Sigma_1 = \begin{bmatrix} 1 & -1.98 \\ -1.98 & 4 \end{bmatrix} \\ \Sigma_2 = \begin{bmatrix} 1 & 1.98 \\ 1.98 & 4 \end{bmatrix} \\ \Sigma_3 = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix} \\ \Sigma_4 = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \end{cases}$$

$$d_1^2 = \begin{bmatrix} x-2 \\ y-3 \end{bmatrix}^T \times \begin{bmatrix} 1 & -1.98 \\ -1.98 & 4 \end{bmatrix}^{-1} \times \begin{bmatrix} x-2 \\ y-3 \end{bmatrix}$$
$$= 50.251x^2 + 49.748xy - 350.251x + 12.562y^2 + 612.562 - 174.874y$$

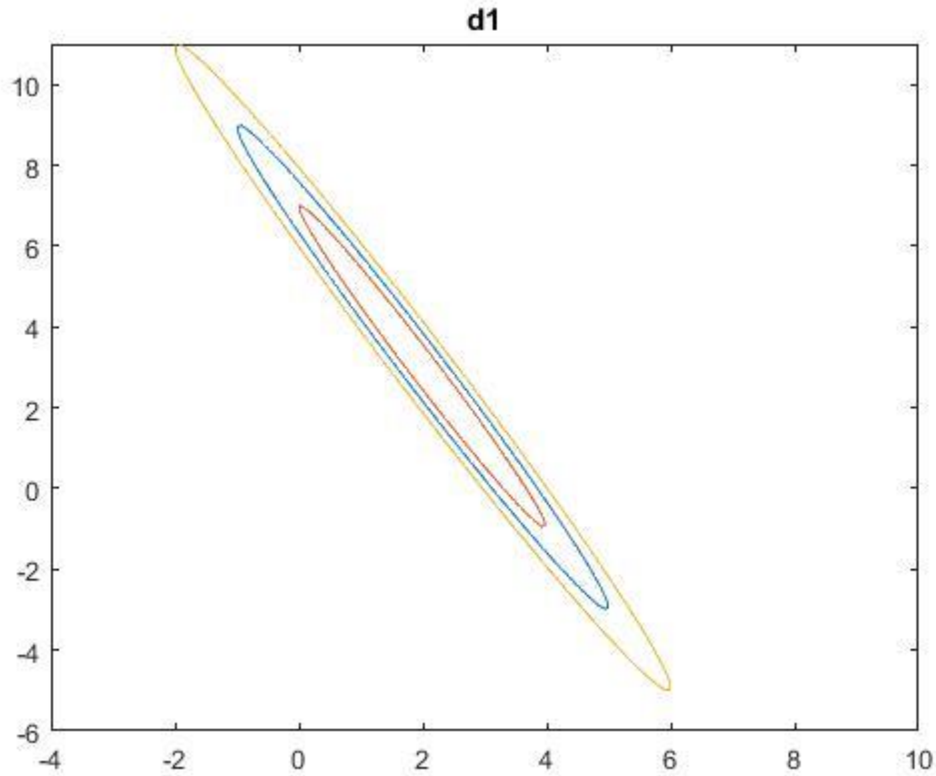
$$d_2^2 = \begin{bmatrix} x-2 \\ y-3 \end{bmatrix}^T \times \begin{bmatrix} 1 & 1.98 \\ 1.98 & 4 \end{bmatrix}^{-1} \times \begin{bmatrix} x-2 \\ y-3 \end{bmatrix}$$
$$= 50.251x^2 - 49.748xy - 51.758x + 12.562y^2 + 15.577 + 24.120y$$

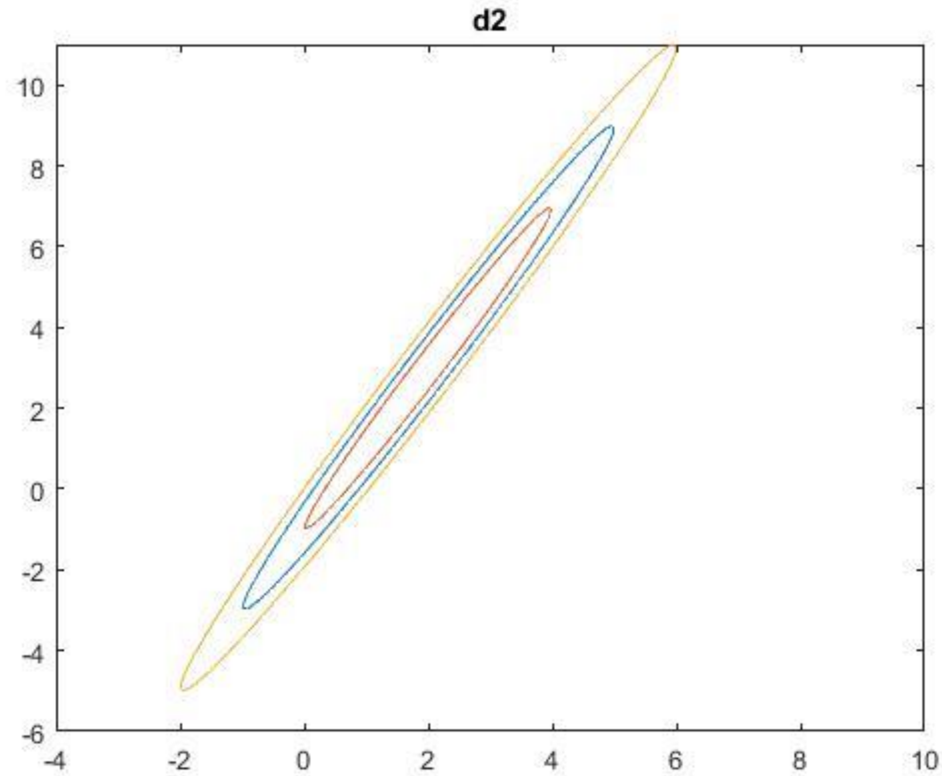
$$d_3^2 = \begin{bmatrix} x-2 \\ y-3 \end{bmatrix}^T \times \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}^{-1} \times \begin{bmatrix} x-2 \\ y-3 \end{bmatrix}$$
$$= 1.333x^2 - 0.666xy - 3.333x + 0.333y^2 + 4.333 - 0.666y$$

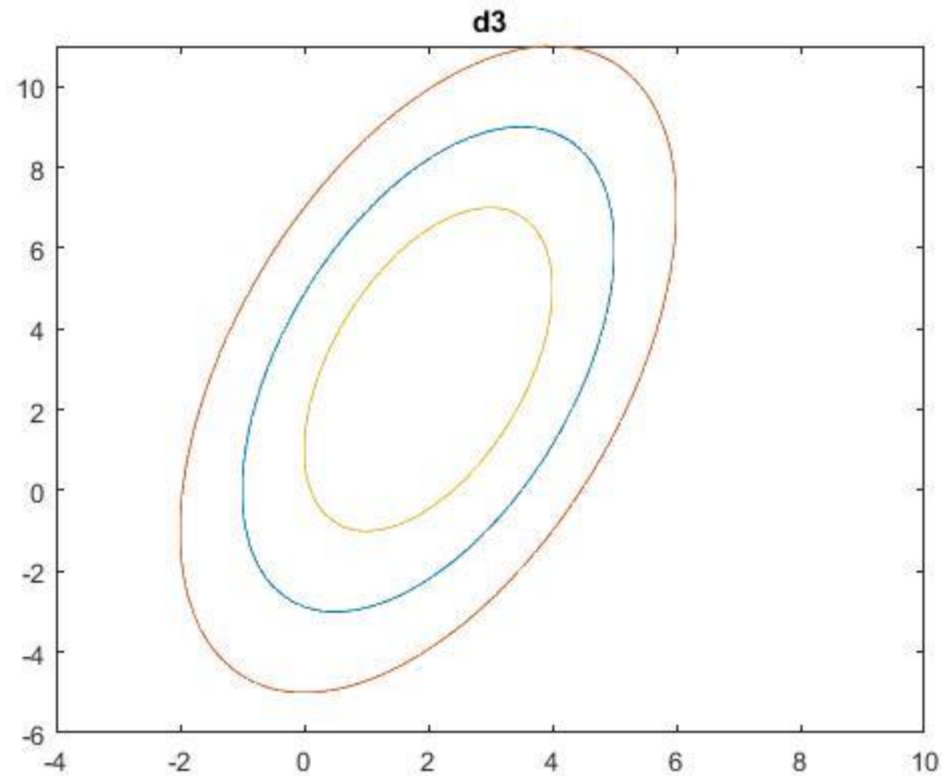
$$d_4^2 = \begin{bmatrix} x-2 \\ y-3 \end{bmatrix}^T \times \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix}^{-1} \times \begin{bmatrix} x-2 \\ y-3 \end{bmatrix}$$
$$= 1.333x^2 + 0.666xy - 7.333x + 0.333y^2 + 12.333 - 3.333y$$

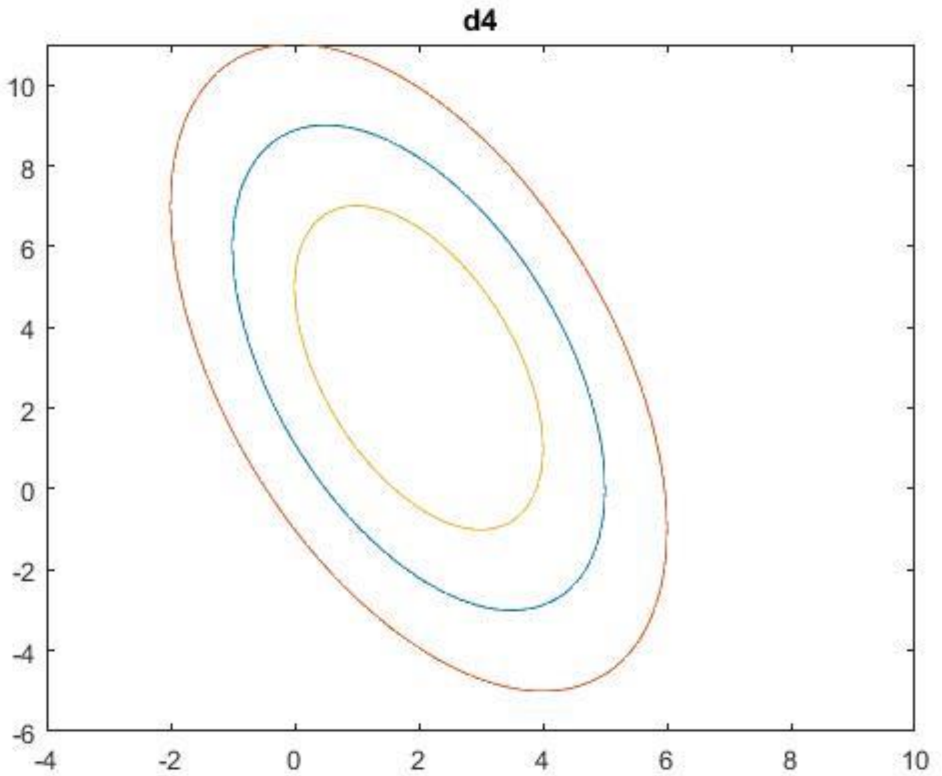
e) اسکریپت e.m از برنامه متلب

تصاویر مربوط به هر معادله قسمت d ( به ازای فواصل ۴و۹و۱۶) در شکل های زیر رسم شده است.




d2
10
8
6
4
2
0
-2
-4
-6
-4
-2
0
2
4
6
8
10


d3
10
8
6
4
2
0
-2
-4
-6
-4
-2
0
2
4
6
8
10




f) اسکریپت b.m از برنامه متلب – بخش f (با کامنت مشخص شده است)

همانطور که مشاهده می‌شود، میانگین و کوواریانس داده‌ها با اعداد ارائه شده توسط سوال اندکی تفاوت دارند. این موضوع به دلیل تصادفی بودن نمونه های تولید شده است.

```
sigma_h =

    2.1187    1.0572
    1.0572    2.8110


mu_h =

    2.0095    0.9626
```

g) اسکریپت b.m از برنامه متلب – بخش g (با کامنت مشخص شده است)

ماتریس A و B Simultaneously diagonalise هستند اگر ضرب آنها جابه‌جا پذیر باشد.(AxB=BxA) ماتریس sigma و sigma_hat به وضوح این شرط را ندارند. به هرحال جهت نشان دادن نتیجه، عمل قطری سازی را با استفاده از ماتریس بردارهای ویژه sigma انجام می‌دهیم:

```
V =

   -0.8507    0.5257
    0.5257    0.8507
```

حال برای قطری کردن، عمل $v^{-1} \times \sigma \times v$ را روی دو ماتریس انجام میدهیم. همانطور که در خروجی مشاهده می کنید، فقط یکی از این دو ماتریس قطری شد. اگرچه به علت نزدیک بودن دو ماتریس، اعداد قطر فرعی ماتریس sigma_hat نزدیک به صفر هستند.

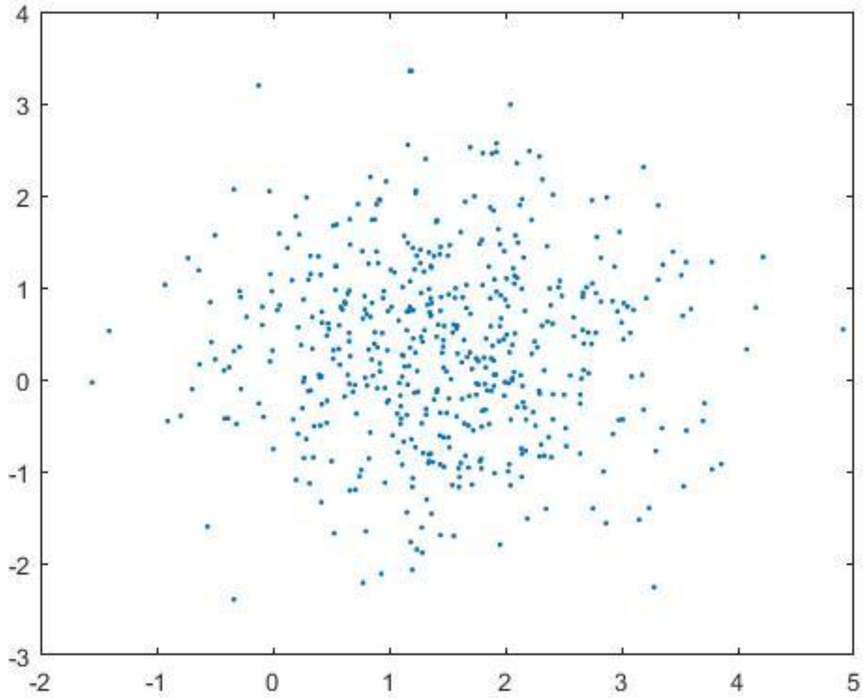
```
ans =

    1.3820   -0.0000
   -0.0000    3.6180


ans =

    1.4039    0.0029
    0.0029    3.0781
```

h) اسکریپت b.m از برنامه متلب – بخش h (با کامنت مشخص شده است)

تبدیل مورد نظر به شکل زیر است:

$$x' = x(\sigma^{\frac{1}{2}})^{-1}$$

بعد از اعمال تبدیل فوق روی داده های بخش b خروجی به شکل زیر است.




همانطور که مشاهده می شود، همبستگی بین x1 و x2 از بین رفته و توزیع داده ها در دو جهت یکسان شده است. کوواریانس داده های جدید در خروجی زیر مشاهده می شود:

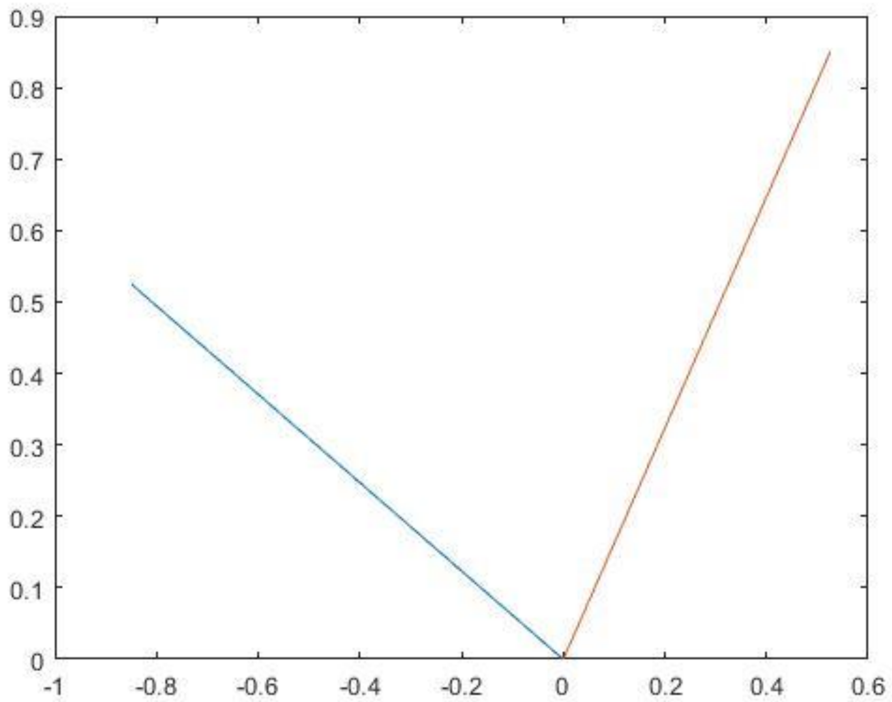
```
part_h_cov =

    1.0000   -0.0000
   -0.0000    1.0000
```

i) اسکریپت b.m از برنامه متلب – بخش i (با کامنت مشخص شده است)

در خروجی زیر ماتریس بردارهای ویژه و ماتریس قطری مقدارهای ویژه ارائه شده است.

```
V =

   -0.8507    0.5257
    0.5257    0.8507


D =

    1.3820         0
         0    3.6180
```

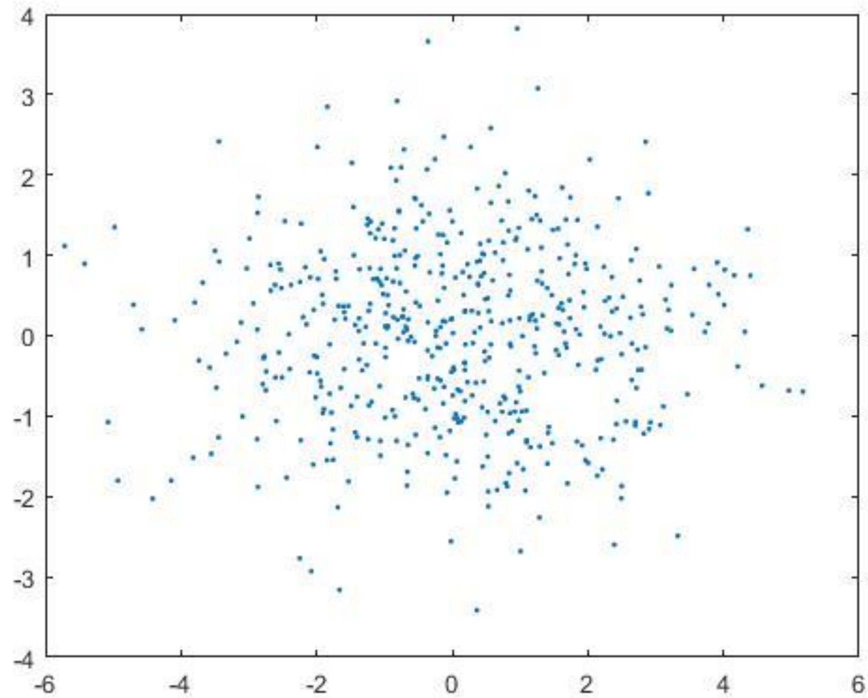
در تصویر زیر، بردارهای ویژه ماتریس رسم شده اند.(این دو بردار بر هم عمود هستند)




بردارهای ویژه، جهت توزیع داده ها را در فضا نشان میدهند. به بیان بهتر، این بردارها، جهت قطر های اصلی بیضی رسم شده در قسمتb هستند. از طرف دیگر، مقادیر ویژه، مقدار توزیع داده ها در دو جهت فوق را نشان می‌دهد.

j) اسکریپت b.m از برنامه متلب – بخش j (با کامنت مشخص شده است)

بعد از اعمال تبدیل موردنظر، داده ها را رسم می کنیم.



باتوجه به شکل، در مقایسه با قسمت b ، میانگین داده ها در مبدا مختصات قرارگرفته و همبستگی(correlation) بین داده ها از بین رفته است.

k) اسکریپت b.m از برنامه متلب – بخش k (با کامنت مشخص شده است)

```
cov_n =

    1.0000    0.0000
    0.0000    1.0000


V_N =

     0     1
     1     0


D_N =

    1.0000         0
         0    1.0000
```

نتایج به دست آمده مطابق انتظار است و با توضیحات داده شده در قسمت i (درباره بردارها و مقدارهای ویژه) مطابقت دارد. بردارهای ویژه جهت توزیع داده ها و مقدارهای ویژه مقدار توزیع داده ها را در آن دو جهت نشان میدهد. با توجه به اینکه کوواریانس I می باشد. بردارهای ویژه، به دوجهت اصلی اشاره می کنند و مقادیر ویژه هر دو، ۱ هستند.

# ٦)

a) این قضیه بیان می کند که مجموع تعداد زیادی متغیر تصادفی مستقل، که هر یک میانگین و واریانس خود را دارند، به طور تقریبی دارای توزیع نرمال خواهد بود. بعضی از مهمترین کاربردهای این قضیه به شرح زیر است:

الف) اگر توزیع نرمالی داشته باشیم که پارامترهای آن مشخص نباشد. با نمونه گیری های متعدد و مستقل از جامعه آماری می توان طبق قضیه حدمرکزی پارامترهای توزیع را تخمین زد.

ب) اگر توزیعی مجهول و غیر نرمال داشته باشیم و n بار از این توزیع تعدادی مساوی نمونه برداریم. طبق قضیه حد مرکزی میانگین نمونه ها دارای توزیع نرمال خواهد بود و بدین ترتیب می توان میانگین و واریانس توزیع مجهول را یافت، حتی اگر توزیع مجهول نرمال نباشد.

b) هر متغیری که قابل اندازه گیری باشد یک measurement است در صورتی که یک feature ممکن است قابل اندازه گیری باشد یا نباشد.

c) عناصر ماتریس کوواریانس از رابطه $\sigma(x,y)=E[(x-E(x))(y-E(y))]$ بدست می آیند. جابه‌جا شدن متغیرها تاثیری در جواب این رابطه ندارد؛ بنابراین ماتریس کوواریانس همیشه قطری است.

d) با توجه به رابطه $(A-\gamma I)v=0$ که در تمرین ۴ بررسی شد، مقدار ویژه صفر نشان میدهد که دترمینان ماتریس A صفر است و یا به بیان دیگر ماتریس A معکوس پذیر نیست.

c) این تبدیل داده ها را مستقل و واریانس آن ها را برابر با یک می کند. استقلال داده ها در مدلهای احتمالاتی مفید است زیرا توابع احتمال توام را ساده تر میکند. همچنین برابر شدن واریانس ها باعث می شود اهمیت همه ابعاد یکسان شود. این موضوع در الگوریتم هایی مانند گرادیان نزولی مهم می باشد.