بسم الله الرحمن الرحیم

گزارش پروژه چهارم

جبر خطی و کاربرد ها
دکتر ناظرفرد

فاطمه قدرتی ۴۰۰۳۱۰۳۸

تاریخ تحویل: ۱۴۰۲/۱۰/۱۴

قسمت اول: درک ماتریس کوواریانس در توزیع گاوسی دو بعدی

گام اول: تشکیل توزیع گاوسی دو بعدی حول محور های X و Y

```
# Step 1: Generate 2D Gaussian distribution along X and Y axes independently
mean = [0, 0]
covariance_matrix = [[1, 0], [0, 1]]
num_points = 10000
```

تابع multivariate_normal نمونه های تصادفی را از یک توزیع نرمال می کشد.
چنین توزیعی با میانگین و ماتریس کوواریانس آن مشخص می شود که مشابه میانگین (میانگین یا "مرکز") و واریانس (انحراف استاندارد مربع یا "عرض") توزیع نرمال یک بعدی است.

```
# Generate random points with Gaussian distribution for both axes
x_points, y_points = np.random.multivariate_normal(mean, covariance_matrix, num_points).T
```

گام دوم: با توجه به مجموعه داده تولید شده (نمونه های تصادفی) ماتریس کوواریانس را محاسبه می کنیم.

```
# Step 2: Calculate the covariance matrix for the generated dataset
dataset = np.vstack((x_points, y_points)).T
calculated_covariance = np.cov(dataset, rowvar=False)
```

گام سوم: نمایش نقاط تشکیل شده (در فایل پروژه)

اهمیت ماتریس کوواریانس:

ماتریس کوواریانس یک ماتریس متقارن است که کوواریانس بین ابعاد مختلف یک مجموعه داده را خلاصه می کند. در زمینه توزیع گاوسی دوبعدی، اطلاعاتی در مورد روابط بین دو متغیر (X و Y) ارائه می‌کند.

عناصر قطری: عناصر قطری ماتریس کوواریانس نشان دهنده واریانس هر متغیر است. مقادیر بزرگتر نشان دهنده گسترش یا تنوع بیشتر در متغیر مربوطه است.

عناصر خارج از قطر: عناصر خارج از قطر نشان دهنده کوواریانس بین متغیر ها هستند. مقادیر مثبت دلالت بر یک همبستگی مثبت دارند (افزایش یک متغیر با افزایش متغیر دیگر همراه است)، در حالی که مقادیر منفی دلالت بر یک همبستگی منفی دارند (افزایش یک متغیر با کاهش متغیر دیگر همراه است).

۲. عناصر مورب و خارج از مورب:

عناصر مورب (واریانس): در توزیع گاوسی دوبعدی، عناصر مورب نشان دهنده واریانس متغیر های منفرد (X و Y) هستند. واریانس های بزرگتر نشان دهنده گسترش یا پراکندگی بیشتر در امتداد هر محور است.

عنصر خارج از قطر (کوواریانس): عنصر خارج از مورب نشان دهنده کوواریانس بین X و Y است. مقدار مثبت نشان می دهد که X و Y تمایل به افزایش یا کاهش با هم دارند، در حالی که مقدار منفی نشان دهنده یک رابطه معکوس است.

قسمت دوم: تبدیل توزیع گاوسی با استفاده از ماتریس [[2,0],[0,3]]

گام چهارم: تشکیل ماتریس تبدیل

```
# Step 4: Generate transformation Matrix
transformation_matrix1 = np.array([[2, 0], [0, 3]])
transformed_dataset1 = np.dot(transformation_matrix1, dataset)
```

گام پنجم: محاسبه ماتریس کوواریانس برای مجموعه داده تبدیل شده

```
# Step 5: Calculate the covariance matrix for the transformed dataset
transformed_covariance = np.cov(transformed_dataset1)
```

گام ششم: نمایش نقاط (در فایل پروژه)

۱. اثرات مورد انتظار دگرگونی:

مقیاس بندی: ماتریس تبدیل [[2,0],[0,3]] توزیع اصلی را با ضریب ۲ در امتداد محور X و ضریب ۳ در امتداد محور Y مقیاس می کند. این بدان معنی است که گسترش یا پراکندگی توزیع در امتداد محور Y در مقایسه با محور X افزایش می یابد.

جهت: ماتریس تبدیل هیچ چرخشی ایجاد نمی کند زیرا مورب است. توزیع تراز خود را با محورهای مختصات حفظ خواهد کرد.

۲. تأثیر بر مقیاس بندی و جهت گیری:

پوسته پوسته شدن: ضریب پوسته پوسته شدن بزرگتر در امتداد محور Y منجر به توزیع کشیده در امتداد آن محور می شود که باعث می شود در مقایسه با محور X گسترده تر به نظر برسد.

جهت گیری: از آنجایی که ماتریس تبدیل مورب است، جهت توزیع با محورهای مختصات تراز باقی می ماند. هیچ چرخشی وجود نخواهد داشت.

۳. تأثیر بر ماتریس کوواریانس:

اثر مقیاس بندی: عناصر مورب ماتریس کوواریانس برای مجموعه داده تبدیل شده، واریانس های مقیاس شده را در امتداد محورهای X و Y منعکس می کند.

بدون اثر چرخش: عناصر خارج از مورب ماتریس کوواریانس بدون تغییر باقی می‌مانند زیرا هیچ چرخشی در ماتریس تبدیل وجود ندارد.

مشاهده ماتریس کوواریانس برای مجموعه داده تبدیل شده، بینش هایی را در مورد اینکه چگونه عوامل مقیاس در امتداد هر محور بر گسترش و همبستگی بین متغیرها در توزیع تبدیل شده تأثیر می گذارد، ارائه می دهد.

قسمت سوم: تبدیل بیشتر توزیع گاوسی با استفاده از ماتریس [[2,1],[1,3]]

گام هفتم: تشکیل دومین ماتریس تبدیل

```python
# Strp 7: Generate second transformation matrix
transformation_matrix2 = np.array([[2, 1], [1, 3]])
transformed_dataset2 = np.dot(transformation_matrix2, dataset)
```

گام هشتم: محاسبه ماتریس کوواریانس برای مجموعه داده تبدیل شده

```python
# Step 8: Calculate the covariance matrix for the doubly transformed dataset
doubly_transformed_covariance = np.cov(transformed_dataset2)
```

گام نهم: نمایش نقاط (در فایل پروژه)

۱. اثرات پیش بینی شده تحول جدید:

مقیاس‌بندی و چرخش ترکیبی: ماتریس [[2,1],[1,3]] هم مقیاس‌گذاری و هم چرخش را معرفی می‌کند. به طور همزمان توزیع را کشیده و می چرخاند.

۲. اثرات ماتریس [[2,1],[1,3]]:

مقیاس بندی: مشابه تبدیل اول، این ماتریس توزیع را در امتداد هر دو محور مقیاس می کند. با این حال، عناصر خارج از مورب یک اثر کوواریانس را معرفی می‌کنند که بر همبستگی بین X و Y تأثیر می‌گذارد.

چرخش: عناصر خارج از مورب بر چرخش توزیع تأثیر می گذارند. عناصر خارج از مورب بزرگتر در این ماتریس چرخش قوی تری نسبت به تبدیل قبلی ایجاد می کنند.

۳. تأثیر بر ماتریس کوواریانس:

اثر مقیاس‌بندی: عناصر مورب ماتریس کوواریانس برای مجموعه داده‌های تغییر شکل مضاعف، واریانس‌های مقیاس‌شده را در امتداد محورهای جدید منعکس می‌کنند.

اثر چرخش: عناصر خارج از مورب ماتریس کوواریانس اکنون اثر چرخش را به تصویر می‌کشند که نشان‌دهنده کوواریانس بین محورهای X و Y جدید است.

قسمت چهارم: تجسم بردارهای ویژه و مقادیر ویژه ماتریس کوواریانس

گام دهم: به دست آوردن مقادیر ویژه و بردار های ویژه ماتریس کوواریانس تبدیل شده

```python
# Step 10: Get eigenvalues of covariance of transformed dataset
eigenvalues, eigenvectors = np.linalg.eig(doubly_transformed_covariance)
```

گام یازدهم: نمایش بردار های ویژه (در فایل پروژه)

۱. اهمیت بردارهای ویژه و مقادیر ویژه:

بردارهای ویژه: بردارهای ویژه جهت حداکثر واریانس در داده ها را نشان می دهند. آنها محورهای اصلی ماتریس کوواریانس را ارائه می دهند.

مقادیر ویژه: مقادیر ویژه مقدار واریانس در جهت هر بردار ویژه را کمیت می کنند. مقادیر ویژه بزرگتر نشان دهنده واریانس بالاتر در امتداد بردار ویژه مربوطه است.

۲. رابطه با گسترش و جهت گیری:

Spread: طول هر بردار ویژه (مقیاس شده با مقدار ویژه آن) نشان دهنده گسترش توزیع در آن جهت است. طول های بزرگتر نشان دهنده گسترش بیشتر است.

جهت گیری: جهت بردارهای ویژه جهت حداکثر واریانس را تعیین می کند. جهت بردار ویژه اول مربوط به محور تغییرات اولیه است و بردار ویژه دوم نشان دهنده محور ثانویه است.

مشاهده بردارهای ویژه به‌عنوان فلش در نمودار، که با مقادیر ویژه متناظر آن‌ها مقیاس‌بندی شده‌اند، نمایشی بصری از نحوه همسویی توزیع و جهت‌گیری با این جهت‌های اصلی ارائه می‌دهد.

قسمت پنجم: طرح ریزی بر روی بردار ویژه غالب

گام دوازدهم: طرح ریزی روی بردار ویژه غالب

```
# Step 12: Project onto the dominant eigenvector
dominant_eigenvector = eigenvectors[:, np.argmax(eigenvalues)]
projected_points = np.dot(dominant_eigenvector, transformed_dataset2)
```

به عبارت ساده تر، این کد پیش بینی اسکالر هر نقطه در مجموعه داده تبدیل شده را بر روی بردار ویژه غالب محاسبه می کند. نتیجه (نقاط_پیش‌بینی شده) نمایشی یک‌بعدی از مجموعه داده‌ها را در جهت حداکثر واریانس ارائه می‌کند و مهم‌ترین اطلاعات را در فضایی با ابعاد پایین‌تر ثبت می‌کند.

این نوع پیش بینی اغلب در تجزیه و تحلیل مؤلفه اصلی (PCA) استفاده می شود، جایی که بردارهای ویژه غالب و مقادیر ویژه متناظر با آنها برای کاهش ابعاد داده ها و در عین حال حفظ مهم ترین ویژگی ها استفاده می شود.

گام سیزدهم: نشان دادن نقاط (در فایل پروژه)

۱. اهمیت طرح ریزی بر روی بردار ویژه غالب:

حداکثر واریانس گرفتن: طرح داده ها بر روی بردار ویژه غالب، داده ها را در جهت حداکثر واریانس تراز می کند. که به ما به گرفتن مهم ترین اطلاعات در مجموعه داده کمک می کند.

۲. تأثیر بر توزیع اصلی:

ثبت واریانس: نقاط پیش‌بینی‌شده حداکثر واریانس موجود در توزیع اصلی را حفظ می‌کنند و به طور موثر جهت اصلی تغییرات مجموعه داده را خلاصه می‌کنند.

کاهش ابعاد: با فرافکنی بر روی بردار ویژه غالب، ابعاد داده ها را کاهش می دهیم و در عین حال حیاتی ترین اطلاعات را در مورد تغییرپذیری آن حفظ می کنیم

مشاهده نقاط اصلی و پیش بینی شده در طرح نشان می دهد که چگونه طرح ریزی داده ها را در جهت حداکثر واریانس متمرکز می کند و یک نمایش ساده ارائه می دهد که اطلاعات ضروری را در مورد مجموعه داده حفظ می کند.