به نام خدا

دانشگاه تهران

دانشکده مهندسی برق و کامپیوتر

مدل‌های مولد عمیق

تمرین شماره ۱

نام و نام خانوادگی	مهسا ندافی قهنویه
شماره دانشجویی	۸۱۰۱۰۰۴۹۰
تاریخ ارسال گزارش	1401/08/13

سوال 5

سوال امتیازی: چرا فرض non-gaussianity در این سوال در نظر گرفته شده است؟

نشان داده میشود که هنگام کار با داده‌های پیوسته، می‌توان با فرض گاوسی به مزیت مهمی دست یافت. در حالی که رویکرد خطی-گاوسی معمولاً تنها به مجموعه‌ای از مدل‌های ممکن منتهی می‌شود، که در ساختار همبستگی شرطی آنها معادل است، یک تنظیم non-gaussian اجازه می‌دهد تا مدل علی کامل، بدون پارامترهای نامشخص، تخمین زده شود.

ایده اصلی پشت این رویکرد این است که توزیع‌های گوسی برای درک برخی از انواع روابط عللی مناسب نیستند، به ویژه در مواردی که مکانیزم‌های غیرخطی یا غیرگوسی علل در حال انجام هستند.

توزیع‌های گوسی فرض می‌کنند که روابط بین متغیرها خطی هستند. با این حال، بسیاری از روابط عللی در جهان واقعی غیرخطی هستند، به این معنی که رابطه بین علت و اثر توسط یک خط مستقیم به خوبی توصیف نمی‌شود. با استفاده از یک مدل غیرگوسی، می‌توانید روابط عللی پیچیده و غیرخطی را بهتر درک کنید. غیرگوسیت علل مخفی را فاش می‌کند. مدل‌های غیرگوسی می‌توانند به کشف علل مخفی که ممکن است بر روی متغیرهای مشاهده شده تأثیر داشته باشند کمک کنند .در یک مدل خطی غیرگوسی، فرض می‌شود که متغیرهای مشاهده شده تحت تأثیر علل مستقیم و علل مخفی قرار دارند. با تحلیل غیرگوسیت داده‌ها، می‌توان از وجود علل مخفی مطلع شد و روابط عللی آنها را استنباط کرد.

a) Let ˆ β12 be the linear regression coefficient when regressing X2 onto X1 (do not consider intercept term). Report ˆ β12 and plot (X1,X2 − ˆ β12X1).

برای خواندن داده از کد زیر استفاده شده است:

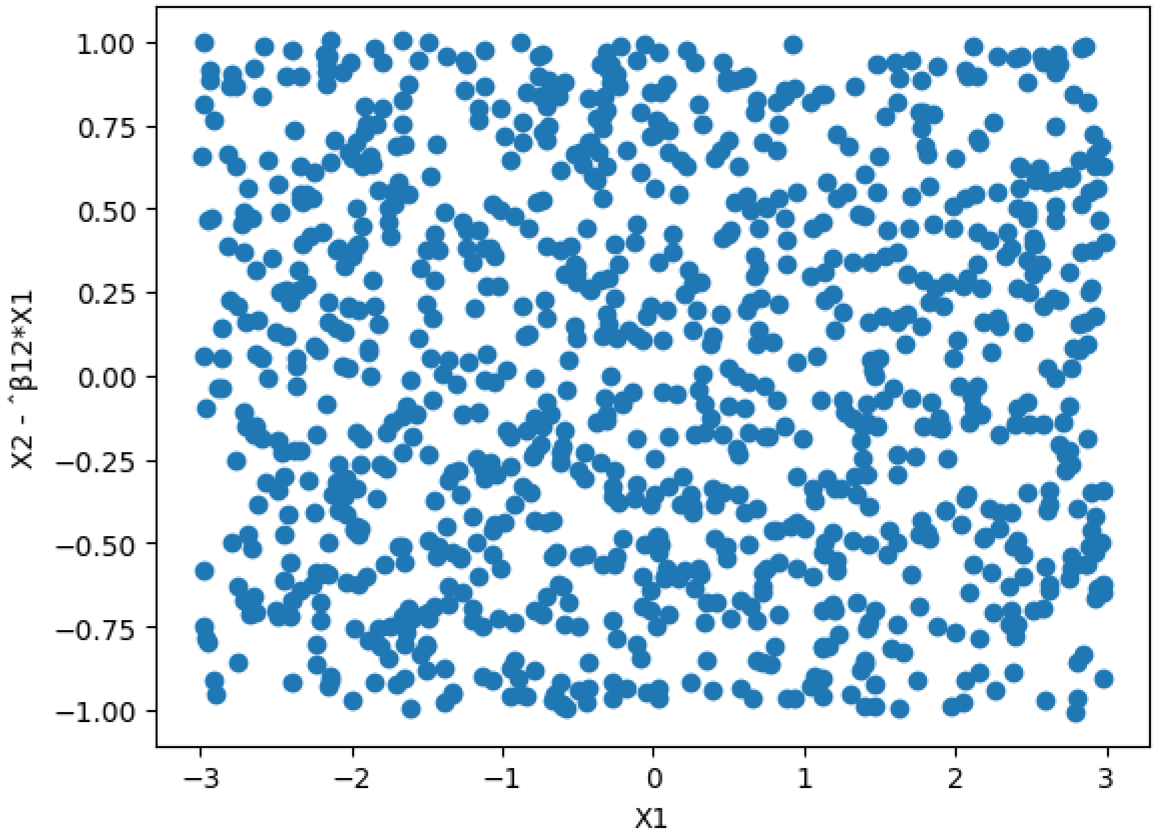
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# Load the data from ngaussian.csv
data = pd.read_csv('ngaussian.csv', header=None,
names=['Combined'])
data[['X1', 'X2']] = data['Combined'].str.split(' ',
expand=True)
data['X1'] = data['X1'].astype(float)
data['X2'] = data['X2'].astype(float)
```

سپس یک مدل رگرسیون خطی بین دو متغیر x1,x2 فیت شده است به طوری که در این بخش x2 متغیر وابسته و x1 متغیر مستقل است.

```
X1 = data['X1'].values.reshape(-1, 1)
X2 = data['X2'].values
# Fit a linear regression model (X2 as the dependent
variable, X1 as the independent variable)
regression_model =
LinearRegression(fit_intercept=False)
regression_model.fit(X1, X2)

# Get the estimated coefficient ^β12
b12_hat = regression_model.coef_[0]

# Calculate the residuals (X2 - ^β12*X1)
residuals = X2 - b12_hat * X1.flatten()

# Print the estimated coefficient ^β12
print("Estimated coefficient ^β12:", b12_hat)

# Plot the residuals
plt.scatter(X1, residuals)
plt.xlabel("X1")
plt.ylabel("c")
plt.show()
```

در نهایت X2 - ^β12*X1 رسم شده و ضریب بتا تخمین زده سده گزارش میشود.

```
Estimated coefficient ^β12: 2.0032651720243653
```



b)Let ˆ β21 be the linear regression coefficient when regressing X1 onto X2 (do not consider intercept term). Report ˆ β21 and plot (X2,X1 − ˆ β21X2).

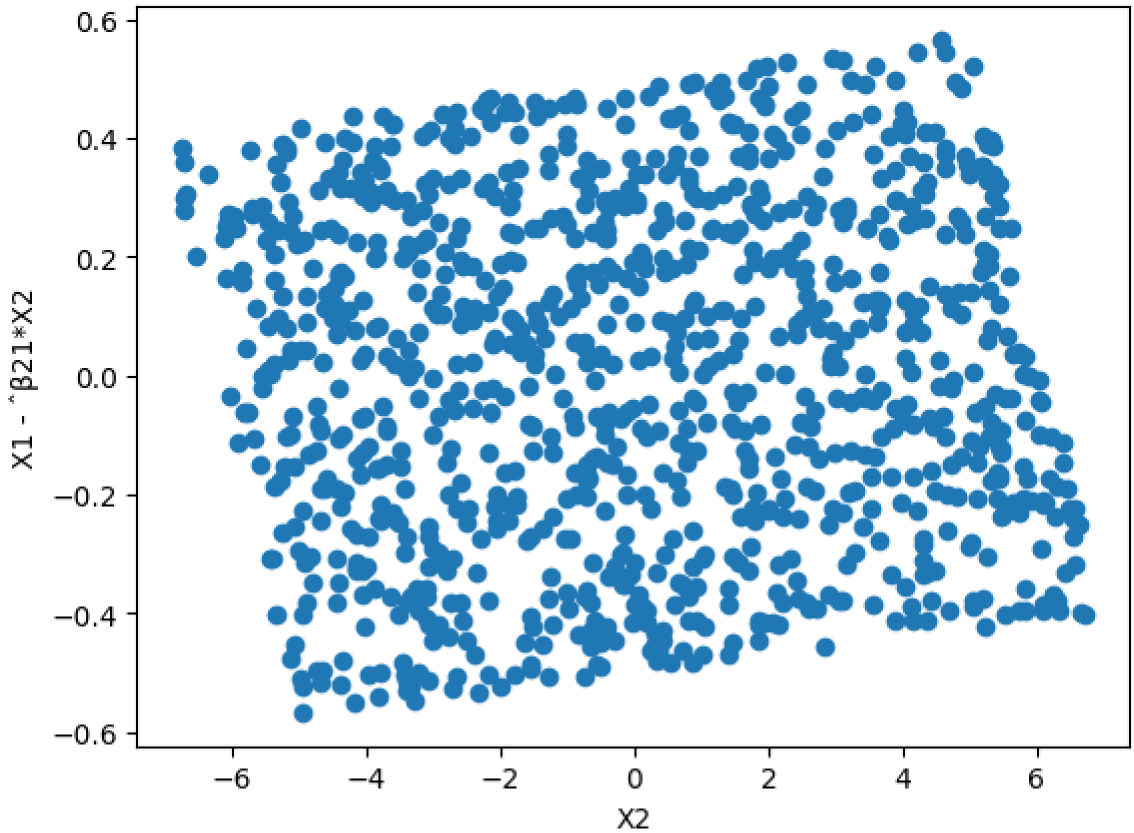
همانند بخش قبل یک مدل رگرسیون خطی بین دو متغیر فیت میکنم اما ایندفعه با فرض اینکه x1 متغیر وابسته و x2 متغیر مستقل است.

```
X2 = data['X2'].values.reshape(-1, 1)
X1 = data['X1'].values

# Fit a linear regression model (X1 as the dependent
variable, X2 as the independent variable)
regression_model =
LinearRegression(fit_intercept=False)
regression_model.fit(X2, X1)

# Get the estimated coefficient ^β21
b21_hat = regression_model.coef_[0]

# Calculate the residuals (X1 - ^β21*X2)
residuals = X1 - b21_hat * X2.flatten()
```

```
# Print the estimated coefficient ^β21
print("Estimated coefficient ^β21:", b21_hat)

# Plot the residuals
plt.scatter(X2, residuals)
plt.xlabel("X2")
plt.ylabel("X1 - ^β21*X2")
plt.show()
```

شکل و ضریب تخمین زده شده:

```
Estimated coefficient ^β21: 0.485379999239924
```



c)

بر این اساس که اگر β21 نزدیک به صفر یا منفی باشد و نمودار X2, X1 - β21X2 یک رابطه با حداقل واریانس را نشان می‌دهد که در مقایسه با نمودار اول شکل بهتری داشته، پس نشان می‌دهد که مدل M2 مجموعه داده را ایجاد کرده است.

به این معنا که X2، cause برای X1 است. X2→ X1

برای اینکه مدل اول درست باشد باید ضریبی نزدیک به یک داشته باشد و شکل آن نشان دهنده فرمولی در خود باشد اما شکل آن کاملا یکنواخت است.