

- HW2
- YASAMAN MIRMOHAMMAD
- 9431022

به نام خدا

گزارش تمرین دوم هوش محاسباتی

- نحوه بازنمایی مسئله دو کلاس ES و کروموزوم، مساله را مدل کرده اند.
- نحوه ترکیب کروموزمها با همدیگر و مقدار احتمال ترکیب و تاثیر آن در سرعت همگرایی. مشاهده شد که هر چه احتمال ترکیب بالاتر برود، سرعت همگرایی بیشتر خواهد شد.

• نحوه جهش و نحوه تعیین گام جهش و تاثیر جهش در سرعت همگرایی.

نحوه ی جهش:

```
chromosome.gene[2] = chromosome.gene[2] * math.exp(-np.random.normal(0, 1) * self.sigma)
chromosome.gene[0] += N01[0] * chromosome.gene[2]
chromosome.gene[1] += N01[1] * chromosome.gene[2]
norm = ((chromosome.gene[0]) ** 2 + (chromosome.gene[1]) ** 2) ** 0.5
chromosome.gene[0] /= norm
chromosome.gene[1] /= norm
self.sigma= np.random.normal(0,0.05);
```

با استفاده از یک تابع گاوسی ، ژن ها را جهش میدهیم و نتیجه ی به دست آمده را نرمال سازی میکنیم.

• نحوه انتخاب والدين و بازماندگان.

```
def choose_parent(self):
    for i in range(1, 4):
        select_index = np.random.randint(0, self.initial_number - 1, self.initial_number)
    for index in select_index:
        self.select_population_set.append(self.initial_population_set[index])
```

```
def choose new generation(self):
    # Todo
    11 11 11
    Use one of the discussed methods in class.
    Q-tournament is suggested !
    :return: Mu selected chromosomes for next cycle
   q_tornoment = 4
   chromosome set = self.select population set
    chromosome_set.extend(self.initial_population_set)
    self.initial population set.clear()
    for i in range(0, self.initial number):
       select index = np.random.randint(0, len(chromosome set) - 1 - i, q tornoment)
        for index in select index:
            select chrom set[j] = chromosome set[index]
        sorted(select_chrom_set, key=lambda x: x.score, reverse=False)
        self.initial_population_set.append(select_chrom_set[0])
        chromosome_set.remove(select_chrom_set[0])
        self.initial population set = sorted(self.initial population set, key=lambda x:
x.score, reverse=False)
   return
```

• تعداد جمعیت و تعداد فرزندان (مقادیر مختلف را امتحان کرده و سرعت همگرایی را مشاهده کرده و نتیجه را ارائه دهید)

تعداد جمعیت و تعداد فرزندان رابطه مستقیمی با سرعت همگرایی دارد چرا که در راستای افزایش تنوع است.

• شرط خاتمه الگوريتم تكاملي.

```
def evolution process(self):
   self.initial_population()
   while (i in range(0, 20)):
       for x in self.initial_population_set:
          print(x.score)
       self.select parent()
       self.mutation()
       self.crossover()
       self.select population()
       print("best chromosome:", self.initial_population_set[0].score,
            statistics.stdev(self.initial population set[0].Z,
statistics.mean(self.initial_population_set[0].Z)))
       print("worst chromosome:", self.initial population set[self.initial number -
1].score)
       print("mean score:", statistics.mean(chrom.score for chrom in
self.initial population set))
   break
```

return self.initial_population_set[0]

اگر به پایان تنوع یا شایستگی برسیم الگوریتم خاتمه یافته است.

توضیحات و نتایج:

PCAیک روش تحلیل داده هاست که درآن، گویی ،محورهای مختصات چرخانده میشوند تا ما از زاویه ی دیگری به توزیع داده هایمان نگاه کنیم.

PCA، تبدیلی در فضای برداری است، که بیشتر برای کاهش ابعاد مجموعه دادهها مورد استفاده قرار می گیرد.

[تحلیل مولفههای اصلی در سال ۱۹۰۱ توسط کارل پیرسون $\frac{1}{2} \frac{1}{4}$ رائه شد. این تحلیل شامل تجزیه مقدارهای ویژه $\frac{1}{2}$ ماتریس کواریانس میباشد].

تعریف:یک تبدیل خطی متعامد است که داده را به دستگاه مختصات جدید میبرد به طوری که بزرگترین واریانس داده بر روی اولین محور مختصات، دومین بزرگترین واریانس بر روی دومین محور مختصات قرار میگیرد و همین طور برای بقیه. تحلیل مولفههای اصلی میتواند برای کاهش ابعاد داده مورد استفاده قرار بگیرد، به این ترتیب مولفههایی از مجموعه داده را که بیشترین تاثیر در واریانس را دارند حفظ میکند

برای ماتریس داده **X،** با میانگین تجربی صفر، که هر سطر یک مجموعه مشاهده و هر ستون دادههای مربوط به یک شاخصه است، تحلیل مولفههای اصلی به صورت زیر تعریف می شود:

به طوری که :

 $Y^T = X^T W = V \Sigma$

به طوری که $\mathbf{V} \mathbf{\Sigma} \mathbf{W}^\mathsf{T}$ تجزیه مقدارهای منفرد ماتریس تج

محدودیت ها:

استفاده از تحلیل مولفههای اصلی منوط به فرض هایی است که در نظر گرفته می شود. از جمله:

- فرض خطی بودن
- ما فرض مي كنيم مجموعه داده تركيب خطى پايههايي خاص است.

- فرض بر این که میانگین و کواریانس از نظر احتمالاتی قابل اتکا هستند.
 - فرض بر این که واریانس شاخصه اصلی داده است.

The covariance between two features x_j and x_k on the population level can be calculated via the following equation:

$$\sigma_{jk} = \frac{1}{n} \sum_{i=1}^{n} (x_j^{(i)} - \mu_j)(x_k^{(i)} - \mu_k)$$

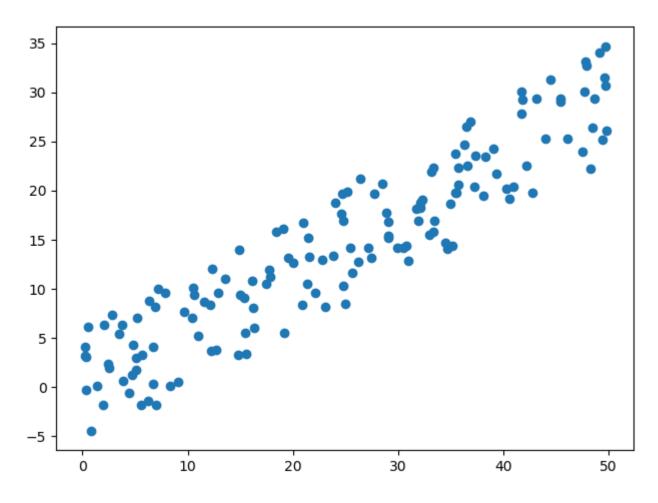
 μ _j and μ _k are the sample means of features j and k.

Want to insert a picture from your files or add a shape, text box, or table? You got it!

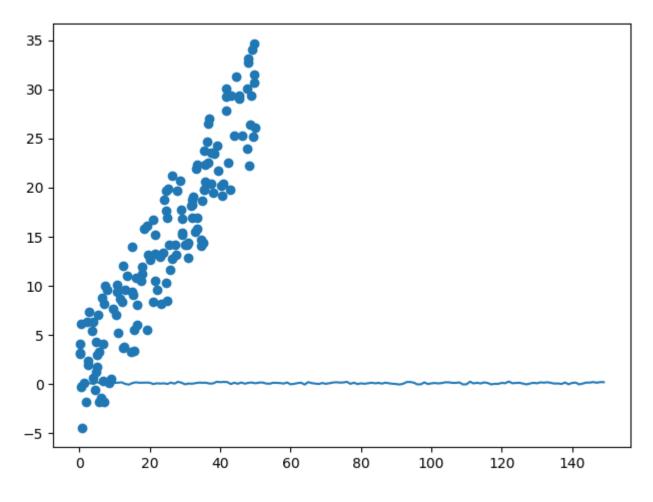
On the Insert tab of the ribbon, just tap the option you need.

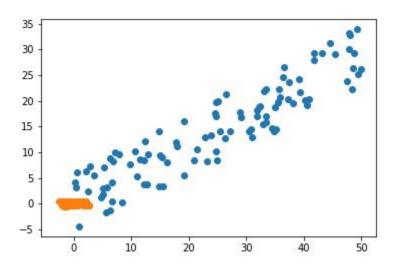
خروجي

داده ها:



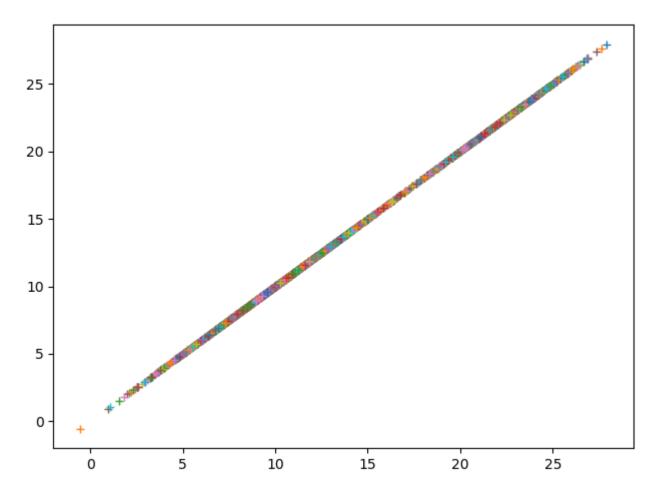
به این شکل مشاهده می شود که داده ها روی یک بعد تصویر شده اند (محور افقی):





نقاط Z تبديل يافته(نارنجي)

خط :pca در محور داده ها:



بدین شکل مشاهده شد که کاهش بعد انجام شده است.