

کاهش ابعاد

سید ناصر رضوی n.razavi@tabrizu.ac.ir

۱۳۹۵

کاهش ابعاد

۲

□ انگیزه.

□ نمایش داده‌ها

□ فشردن سازی داده‌ها

■ کاهش مصرف حافظه

■ استفاده‌ی آسان‌تر از مجموعه داده‌ها

■ کاهش هزینه‌های محاسباتی بسیاری از الگوریتم‌ها

■ حذف نویز و افزایش دقت الگوریتم یادگیری

■ ساده‌تر کردن درک نتایج

□ روش‌ها.

□ تحلیل مؤلفه‌های اصلی

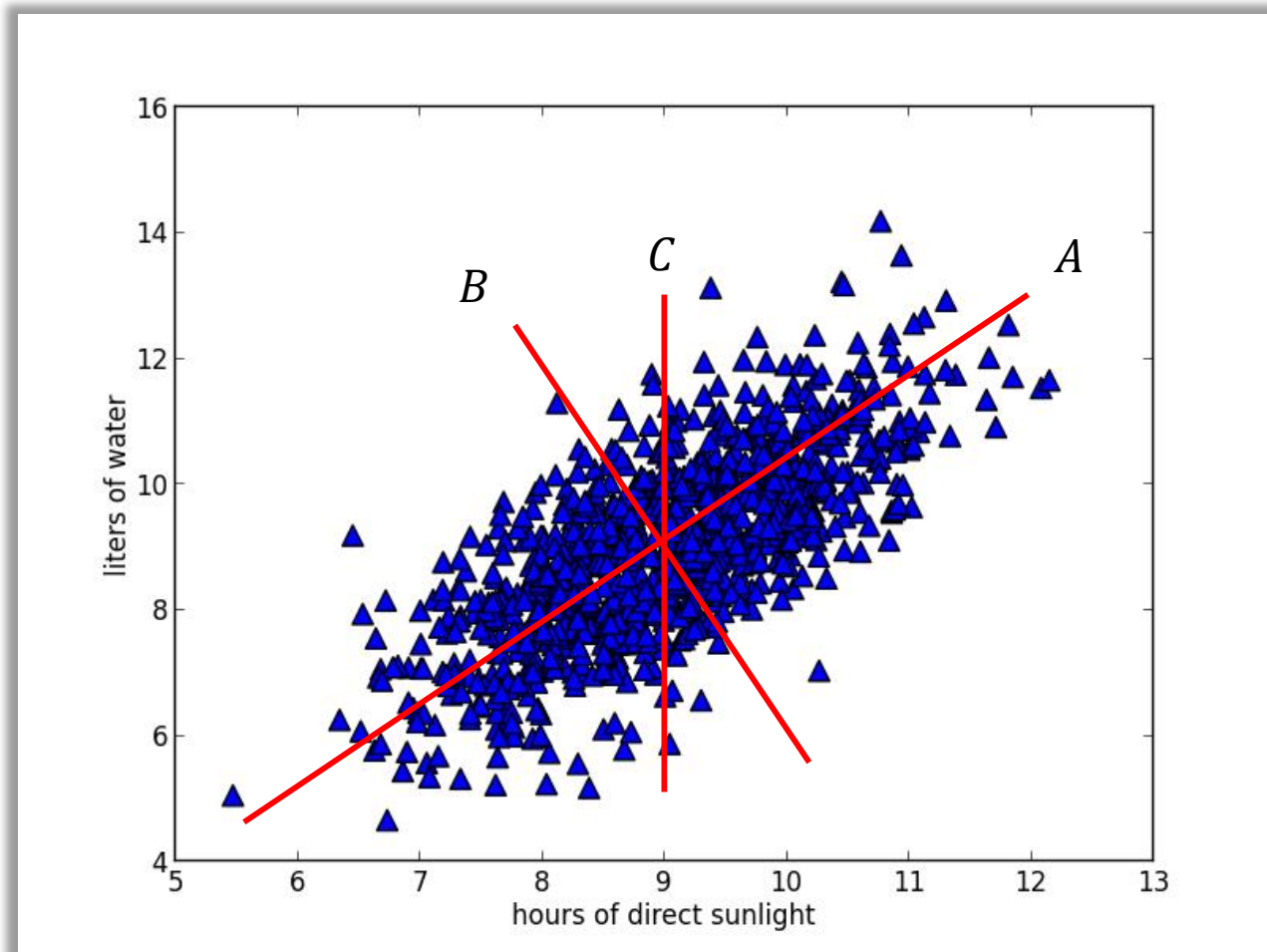
□ تحلیل فاکتورها

□ تحلیل مؤلفه‌های مستقل

تحلیل مؤلفه‌های اصلی (PCA)

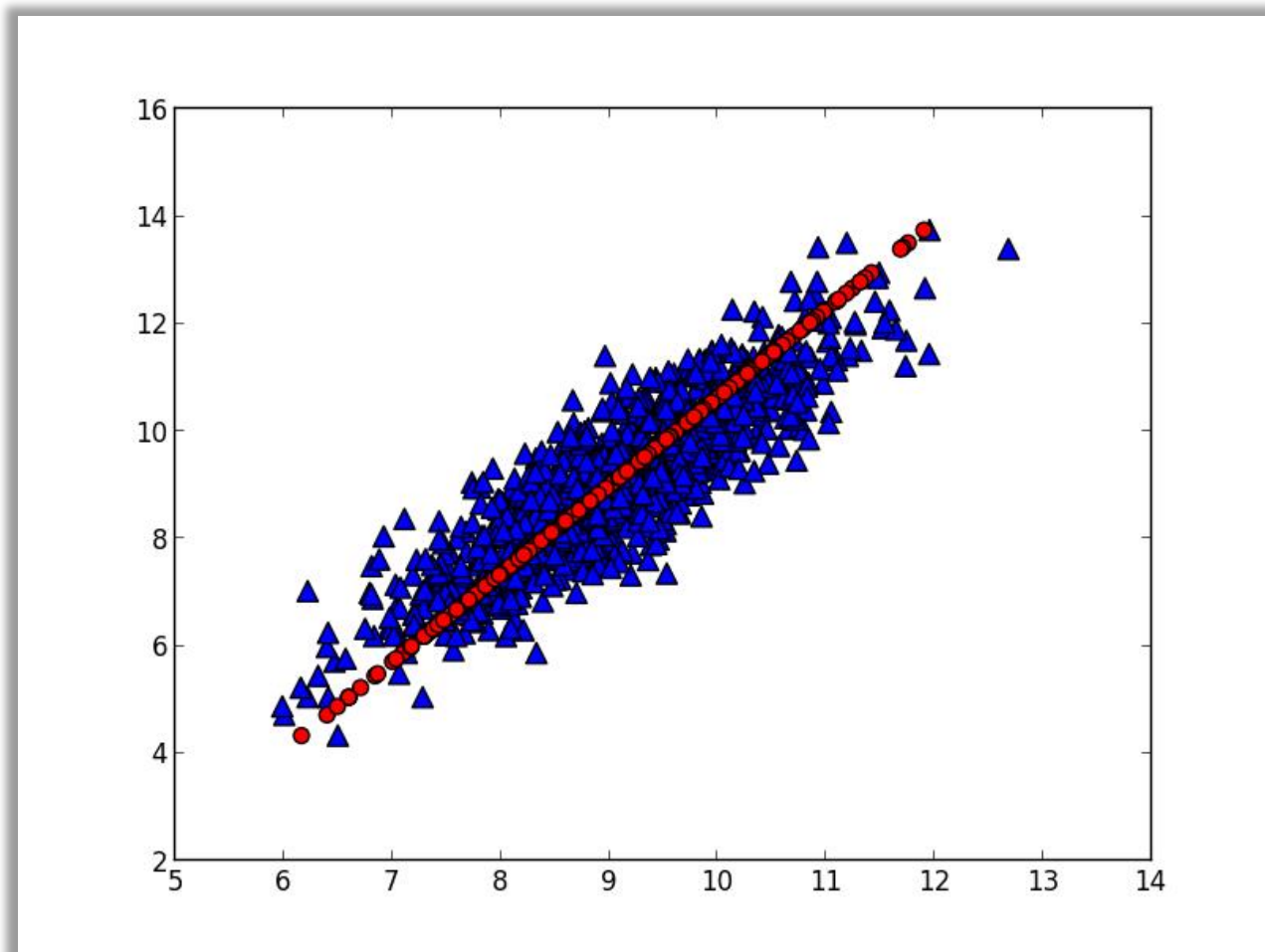
تحلیل مؤلفه‌های اصلی

۴



تحلیل مؤلفه‌های اصلی

۵

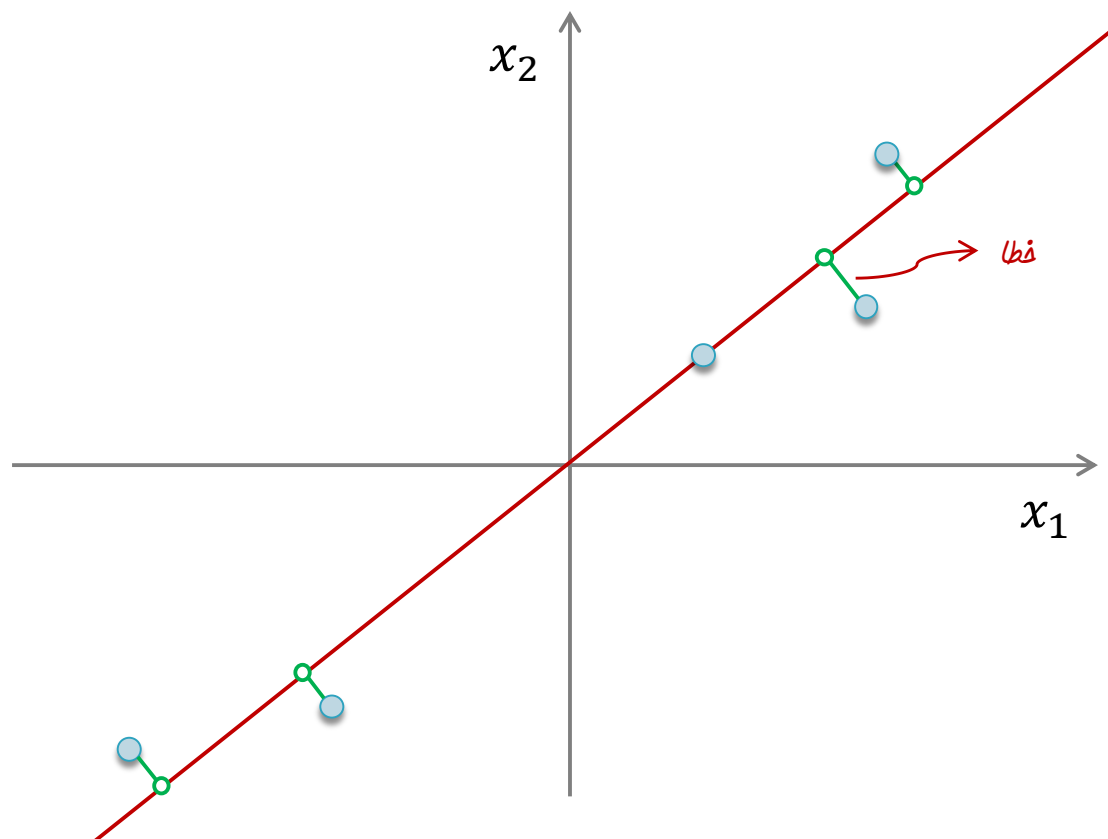


تحلیل مؤلفه‌های اصلی: بیان مسئله

بیان مسئله

۷

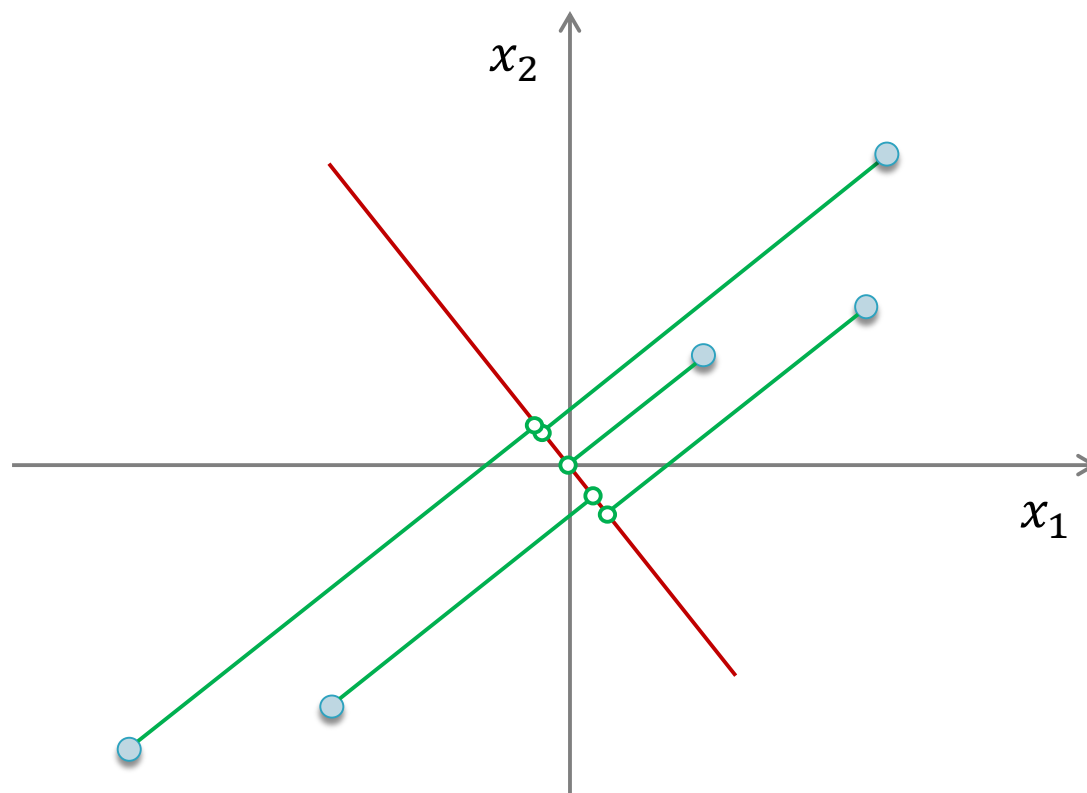
□ هدف. کمینه‌سازی مجموع مربعات خطای تابش



بیان مسئله

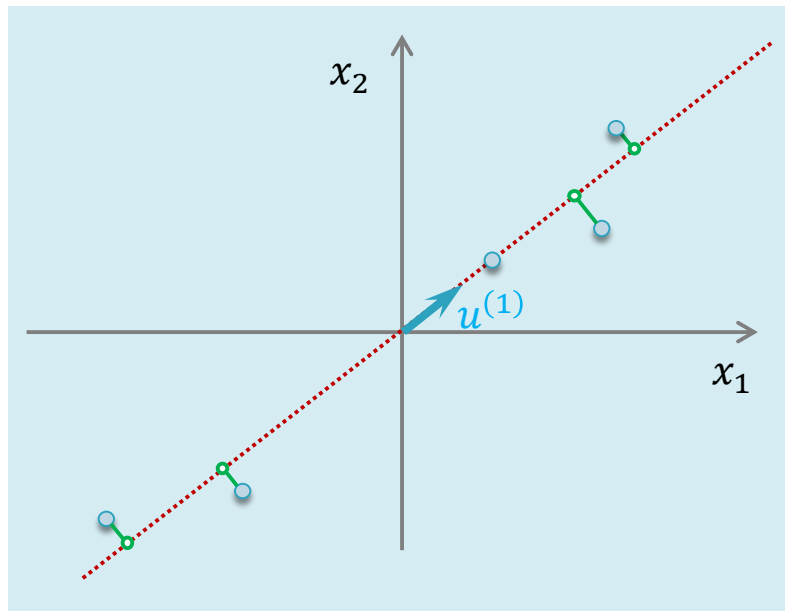
۸

□ هدف. کمینه سازی مجموع مربعات خطای تابش



بیان مسئله

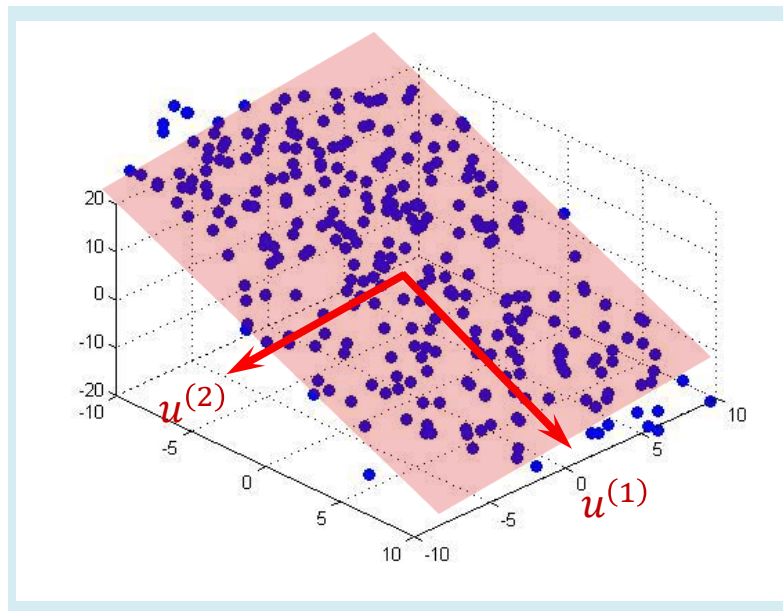
۹



□ کاهش ابعاد از ۲ به ۱. یافتن یک جهت مانند $u^{(1)} \in \mathbb{R}^n$ به طوری که با تصویر کردن نقاط در آن جهت، مجموع مربعات خطا کمینه گردد.

بیان مسئله

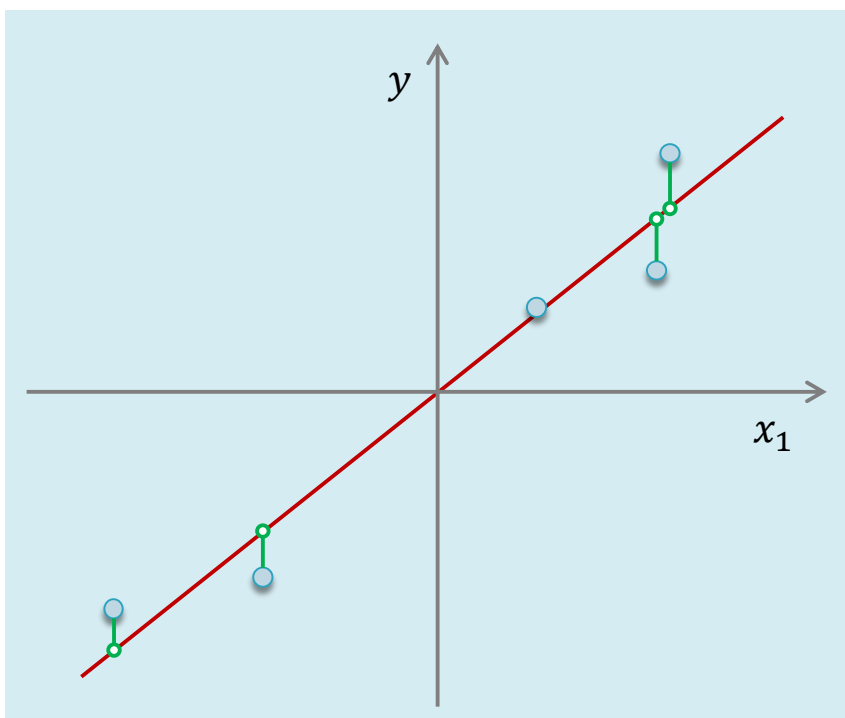
۱۰



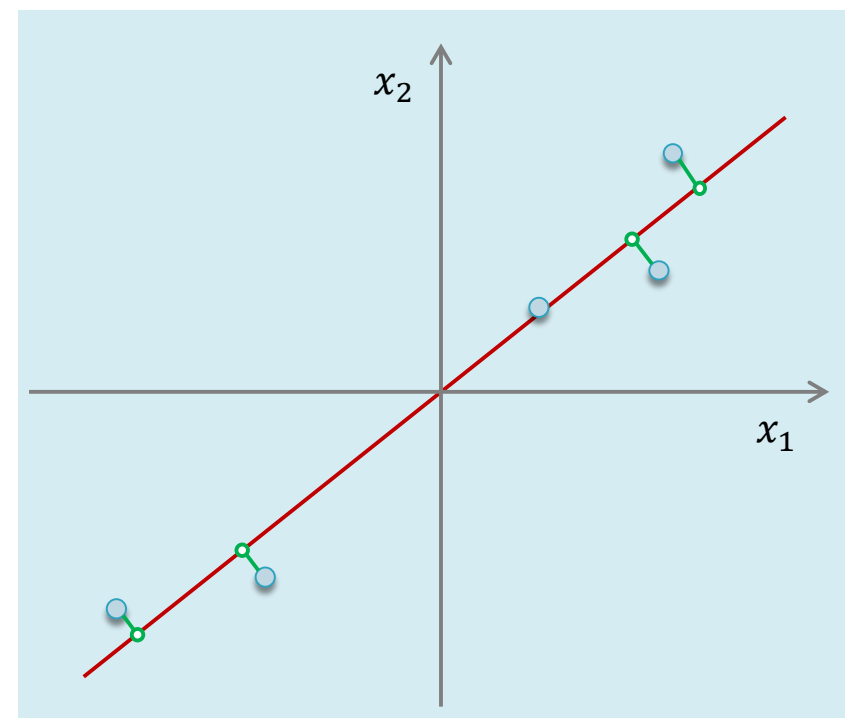
□ کاهش ابعاد از d به k . یافتن k بردار متعامد مانند $u^{(1)}, u^{(2)}, \dots, u^{(k)}$ به طوری که با تصویر کردن نقاط در آن جهتها، مجموع مربعات خطا کمینه گردد.

PCA همان رگرسیون است؟

۱۱



رگرسیون خطی



تحلیل مؤلفه‌های اصلی

الْأَوْرِيْتَه PCA



الگوریتم PCA: پیش پردازش

۱۳

□ مجموعه‌ی آموزشی.

$$\{x^{(1)}, x^{(2)}, x^{(3)}, \dots, x^{(m)}\}, \quad x^{(i)} \in \mathbb{R}^n$$

□ پیش پردازش.

(۱) حذف میانگین.

$$\mu_j = \frac{1}{m} \sum_{i=0}^m x_j^{(i)}, \quad x_j^{(i)} = x_j^{(i)} - \mu_j$$

(۲) مقیاس بندی. [در صورت نیاز]

اگر ویژگی‌های مختلف دارای بازه‌های مختلف هستند (مانند ویژگی‌های اندازه‌ی خانه و تعداد اتاق‌ها)، ویژگی‌ها را به گونه‌ای مقیاس بندی کنید که بازه‌ی ویژگی‌های مختلف تقریباً برابر باشند.

$$x_j^{(i)} = \frac{x_j^{(i)} - \mu_j}{s_j}$$

← انحراف معیار در ویژگی j

الگوریتم PCA

۱۴

□ کاهش ابعاد داده‌ها از d به k .

□ محاسبه‌ی «ماتریس کوواریانس»:

$$\Sigma = \frac{1}{m} X^T X = \frac{1}{m} \sum_{i=1}^n x^{(i)} (x^{(i)})^T$$

□ محاسبه‌ی «بردارهای ویژه» ماتریس کوواریانس:

$$[U, S, V] = \text{svd}(\Sigma)$$

□ انتخاب k بردار اول از ماتریس U :

$$U = \begin{bmatrix} | & | & \dots & | \\ u^{(1)} & u^{(2)} & \dots & u^{(n)} \\ | & | & \dots & | \end{bmatrix}_{n \times n}$$



$$U_{reduced} = \begin{bmatrix} | & | & \dots & | \\ u^{(1)} & u^{(2)} & \dots & u^{(k)} \\ | & | & \dots & | \end{bmatrix}_{n \times k}$$

□ محاسبه‌ی داده‌های جدید با ابعاد k .

$$z_{k \times 1}^{(i)} = \begin{bmatrix} | & | & \dots & | \\ u^{(1)} & u^{(2)} & \dots & u^{(k)} \\ | & | & \dots & | \end{bmatrix}^T \times x_{n \times 1}^{(i)}$$

$$= \begin{bmatrix} - & u^{(1)} & - \\ - & u^{(2)} & - \\ \vdots & \vdots & \vdots \\ - & u^{(k)} & - \end{bmatrix}_{k \times n} \times x_{n \times 1}^{(i)}$$

پیاده سازی در اکتاو

۱۶

□ الگوریتم PCA.

□ پس از حذف میانگین و در صورت نیاز مقیاس بندی.

```
function Z = PCA(X)
```

```
    m = size(X, 1);
```

```
    Sigma = (1/m) * X' * X; ← مماسبه ماتریس کوواریانس
```

```
    [U, S, V] = svd(Sigma); ← مماسبه تجزیه مقادیر منفرد
```

```
    U_reduced = U(:, 1:k); ← انتخاب  $k$  مؤلفه اول
```

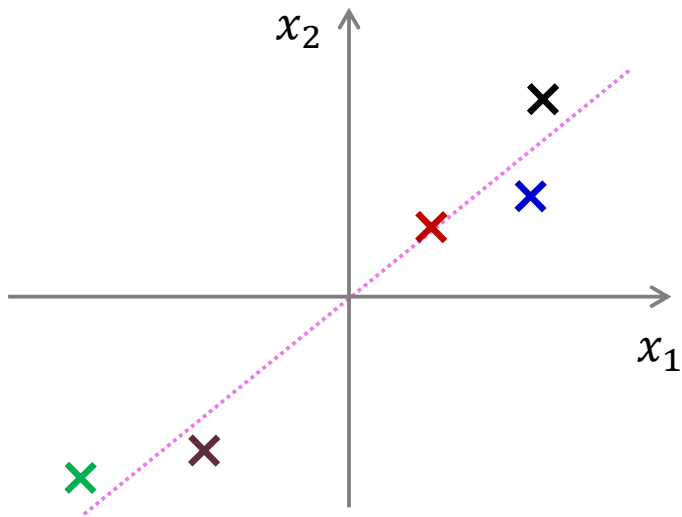
```
    Z = X * U_reduced; ← مماسبه داده های جبر با  $k$  بعد
```

```
end;
```


مثال: کاهش ابعاد

۱۷

$$Z = X \times U_{reduced}$$



داده‌های اولیه (دو بعدی)

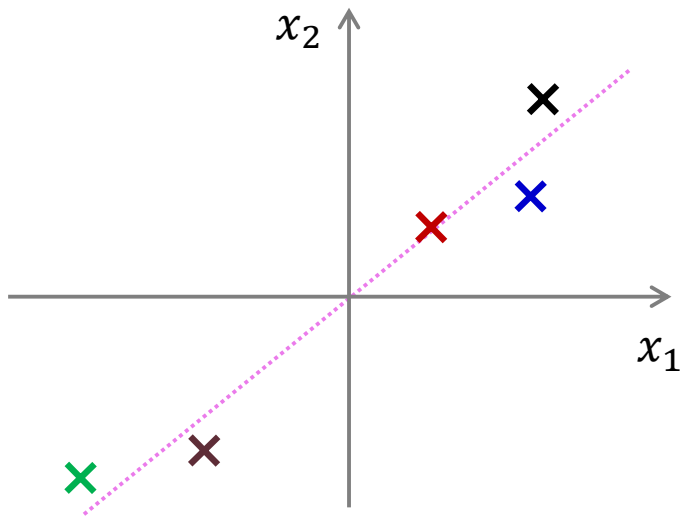


داده‌های جدید (یک بعدی)

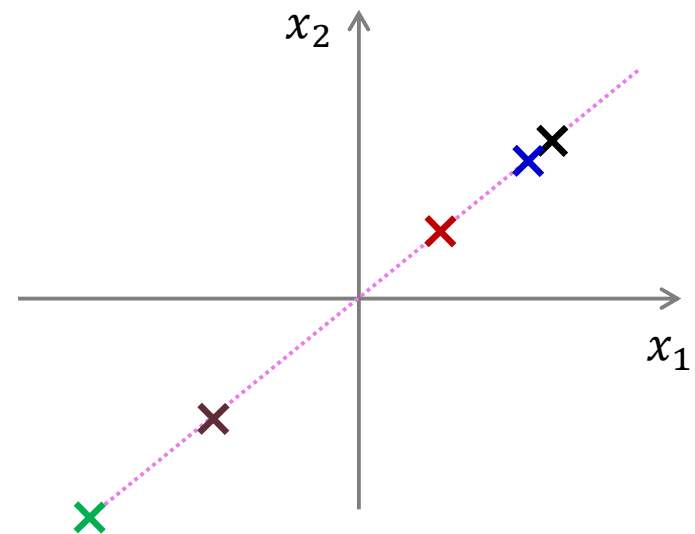
مثال: کاهش ابعاد

۱۸

$$X_{recovered} = Z * U_{reduced}^T + means$$



داده‌های اولیه (دو بعدی)



داده‌های بازسازی شده

الگوریتم PCA: حذف میانگین

۱۹

```
>> X = [1 1 1 0 0; 2 2 2 0 0; 1 1 1 0 0; 5 5 5 0 0;  
        1 1 0 2 2; 0 0 0 3 3; 0 0 0 1 1];
```

```
>> meanVals = mean(X);
```

```
>> X_norm = zeros(size(X));
```

```
>> for i=1:size(X, 1)
```

```
>     X_norm(i, :) = X(i, :) - meanVals;
```

```
>> X_norm
```

```
X_norm =
```

```
-0.42857 -0.42857 -0.28571 -0.85714 -0.85714  
 0.57143  0.57143  0.71429 -0.85714 -0.85714  
-0.42857 -0.42857 -0.28571 -0.85714 -0.85714  
 3.57143  3.57143  3.71429 -0.85714 -0.85714  
-0.42857 -0.42857 -1.28571  1.14286  1.14286  
-1.42857 -1.42857 -1.28571  2.14286  2.14286  
-1.42857 -1.42857 -1.28571  0.14286  0.14286
```

داده‌های اولیه

1	1	1	0	0
2	2	2	0	0
1	1	1	0	0
5	5	5	0	0
1	1	0	2	2
0	0	0	3	3
0	0	0	1	1

```
X_norm = bsxfun(@minus, X, mean(X))
```

الگوریتم PCA: محاسبه بردارهای ویژه

۲۰

```
>> m = size(X, 1);  
>> Sigma = (1/m) * X_nom' * X_norm;  
  
>> [U, S, V] = svd(Sigma);  
>> S  
S =  
Diagonal Matrix  
8.7173e+000      0      0      0      0  
      0 1.5832e+000      0      0      0  
      0      0 6.6876e-002      0      0  
      0      0      0 5.6022e-016      0  
      0      0      0      0 1.4144e-047
```

الگوریتم PCA: کاهش ابعاد

۲۱

```
>> U_red = U(:, 1:3);
```

```
>> Z = X_norm * U_red;
```

```
Z =
```

0.1667425	1.3749474	-0.0091539
-1.4442884	0.7390287	-0.0228180
0.1667425	1.3749474	-0.0091539
-6.2773812	-1.1687275	-0.0638103
1.7595299	-1.1001502	0.5712943
3.3326043	-1.9204675	-0.3520239
2.2960504	0.7004216	-0.1143345

الگوریتم PCA: بازسازی داده های اولیه

۲۲

```
>> X_recovered = Z * U_red';  
>> for i = 1:m  
>     X_recovered(i, :) = X_recovered(i, :) + meanVals;  
>> X_recovered
```

X_recovered =

```
1.00000  1.00000  1.00000  0.00000  0.00000  
2.00000  2.00000  2.00000  0.00000  0.00000  
1.00000  1.00000  1.00000  0.00000  0.00000  
5.00000  5.00000  5.00000 -0.00000 -0.00000  
1.00000  1.00000  0.00000  2.00000  2.00000  
0.00000 -0.00000  0.00000  3.00000  3.00000  
-0.00000 -0.00000 -0.00000  1.00000  1.00000
```

داده های اولیه

1	1	1	0	0
2	2	2	0	0
1	1	1	0	0
5	5	5	0	0
1	1	0	2	2
0	0	0	3	3
0	0	0	1	1

انتخاب تعداد مؤلفه‌های اصلی

انتخاب تعداد مؤلفه‌های اصلی

۲۴

□ میانگین مجموع مربعات خطای تابش.

$$\frac{\frac{1}{m} \sum_{i=1}^m \|x^{(i)} - x_{approx}^{(i)}\|^2}{\frac{1}{m} \sum_{i=1}^m \|x^{(i)}\|^2} \leq 0.01$$

مفظ ۹۹ درصد واریانس

$$\frac{\frac{1}{m} \sum_{i=1}^m \|x^{(i)} - x_{approx}^{(i)}\|^2}{\frac{1}{m} \sum_{i=1}^m \|x^{(i)}\|^2} \leq 0.05$$

مفظ ۹۵ درصد واریانس

$$\frac{\frac{1}{m} \sum_{i=1}^m \|x^{(i)} - x_{approx}^{(i)}\|^2}{\frac{1}{m} \sum_{i=1}^m \|x^{(i)}\|^2} \leq 0.10$$

مفظ ۹۰ درصد واریانس

انتخاب تعداد مؤلفه‌های اصلی

۲۵

یک الگوریتم غیر کارا

$k = 0$

repeat

{

$k = k + 1$

try $PCA(X)$ with k components

compute $U_{reduced}, z^{(1)}, z^{(2)}, \dots, z^{(m)}, x_{approx}^{(1)}, x_{approx}^{(2)}, \dots, x_{approx}^{(m)}$

} until $\frac{\frac{1}{m} \sum_{i=1}^m \|x^{(i)} - x_{approx}^{(i)}\|^2}{\frac{1}{m} \sum_{i=1}^m \|x^{(i)}\|^2} \leq 0.01$

انتخاب تعداد مؤلفه‌های اصلی

۲۶

remove mean from X

$$\text{sigma} = \frac{1}{m} (X' * X);$$

$$[U, S, V] = \text{svd}(\text{sigma});$$

pick smallest value of k for which:

$$\frac{\sum_{i=1}^k s_{ii}}{\sum_{i=1}^n s_{ii}} \geq 0.99$$

$$S = \begin{bmatrix} s_{11} & 0 & \cdots & 0 \\ 0 & s_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_{nn} \end{bmatrix}$$

انتخاب تعداد مؤلفه‌های اصلی

۲۷

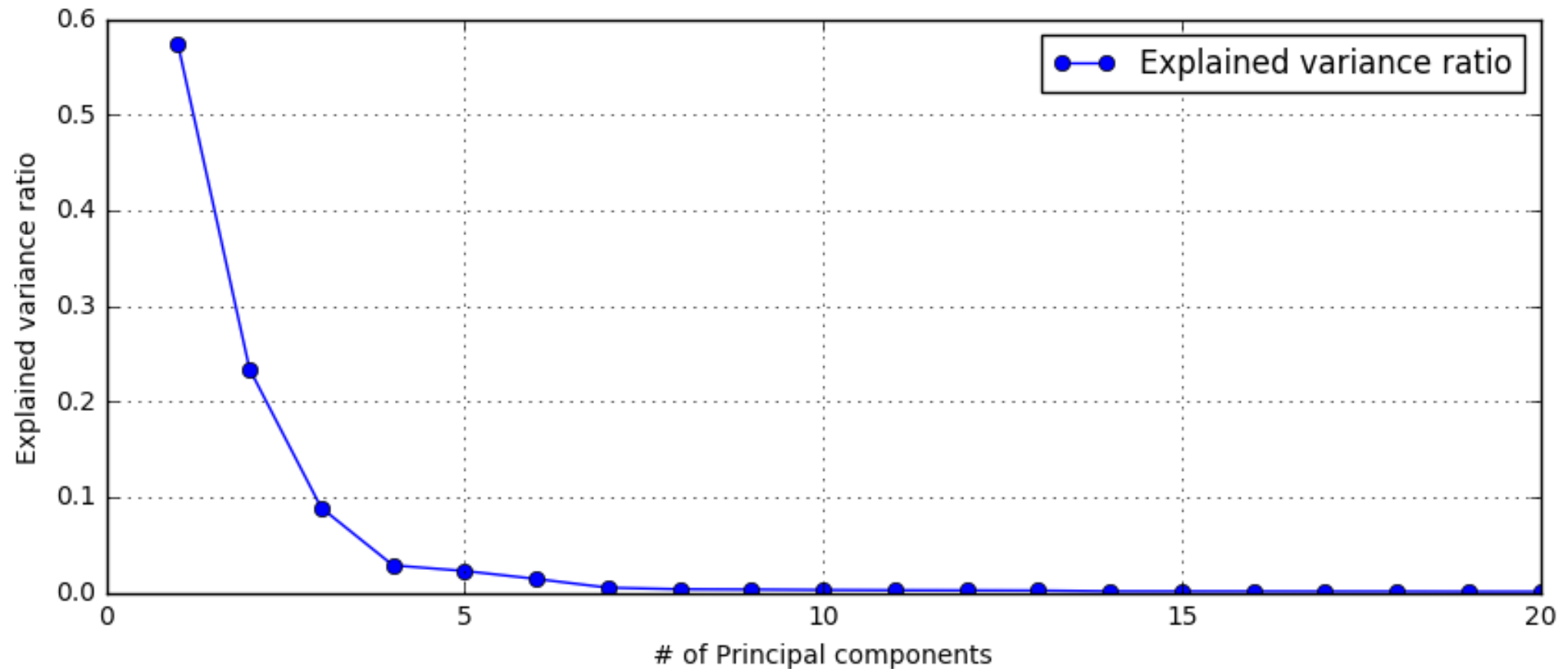
□ مثال. داده‌های مربوط به نیمه هادی‌ها [۵۹۰ ویژگی]

تعداد مؤلفه‌ها	درصد واریانس	درصد تجمعی
۱	۵۹/۲	۵۹/۲
۲	۲۴/۱	۸۳/۴
۳	۹/۲	۹۲/۵
۴	۲/۳	۹۴/۸
۵	۱/۵	۹۶/۳
۶	۰/۵	۹۶/۸
۷	۰/۳	۹۷/۱
۲۰	۰/۰۸	۹۹/۳

<http://archive.ics.uci.edu/ml/datasets/SECOM>

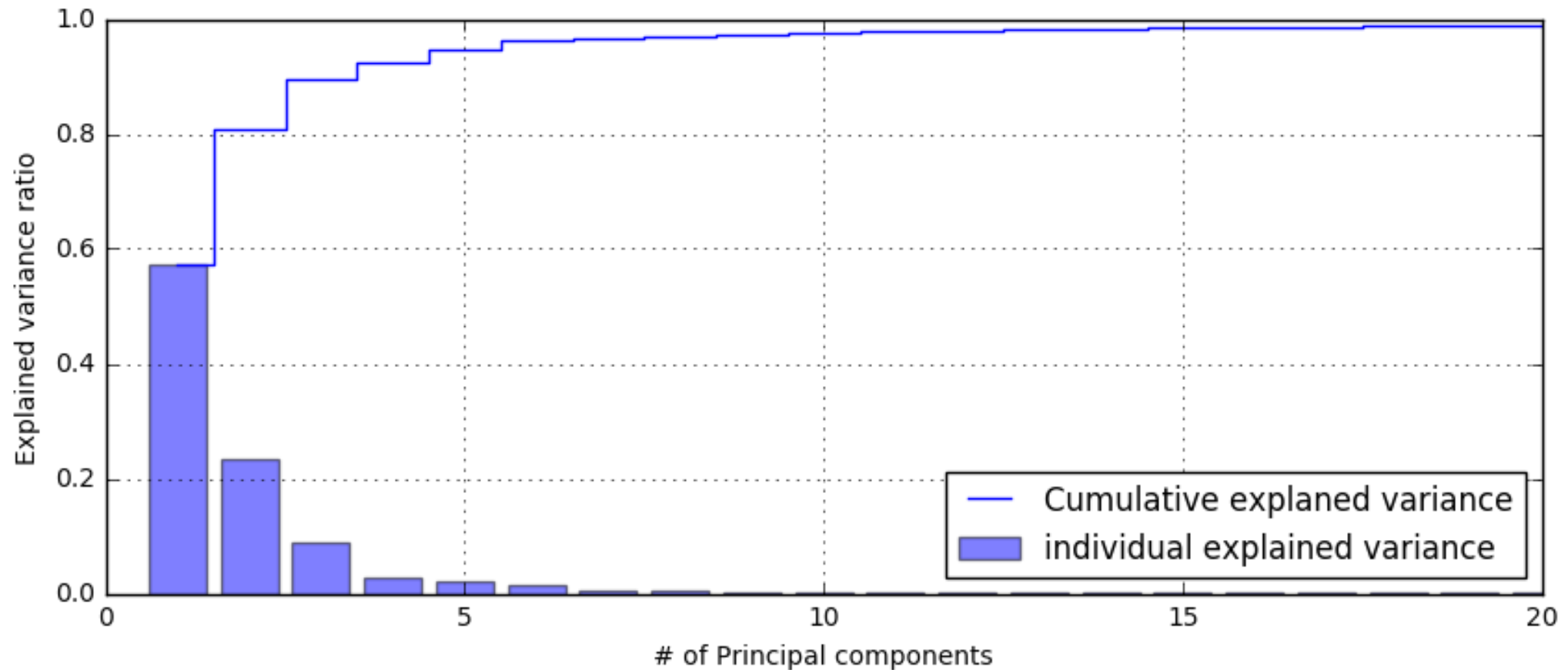
انتخاب تعداد مؤلفه‌های اصلی

۲۸



انتخاب تعداد مؤلفه‌های اصلی

۲۹



توصیه‌های عملی

افزایش سرعت یادگیری نظارت شده

۳۱

□ مجموعه‌ی آموزشی اولیه.

$$\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), (x^{(3)}, y^{(3)}), \dots, (x^{(m)}, y^{(m)})\}$$

□ استخراج ورودی‌ها.

$$\{x^{(1)}, x^{(2)}, x^{(3)}, \dots, x^{(m)}\}, \quad x^{(i)} \in \mathbb{R}^{10000}$$

□ اجرای PCA.

$$\{z^{(1)}, z^{(2)}, z^{(3)}, \dots, z^{(m)}\}, \quad z^{(i)} \in \mathbb{R}^{1000}$$

□ مجموعه‌ی آموزشی جدید.

$$\{(z^{(1)}, y^{(1)}), (z^{(2)}, y^{(2)}), (z^{(3)}, y^{(3)}), \dots, (z^{(m)}, y^{(m)})\}$$

□ **توجه.** نگاشت از فضای d بعدی به فضای k بعدی باید تنها با استفاده از داده‌های موجود در مجموعه‌ی آموزشی (X_{train}) به دست آید.

□ فشرده‌سازی داده‌ها.

- کاهش مصرف حافظه‌ی مورد نیاز برای ذخیره‌سازی داده‌ها
- افزایش سرعت اجرای الگوریتم یادگیری
- انتخاب تعداد مؤلفه‌ها: بر اساس درصد واریانس حفظ شده

□ تصویر نمودن داده‌ها.

- انتخاب تعداد مؤلفه‌ها: $k = 2$ یا $k = 3$

استفاده نادرست از PCA: مقابله با بیش‌برازش

۳۳

□ روش نادرست.

□ استفاده از $z^{(i)}$ به جای $x^{(i)}$ باعث کاهش تعداد ویژگی‌ها از n به k می‌شود؛

□ در نتیجه، با داشتن ویژگی‌های کمتر، احتمال بیش‌برازش کاهش می‌یابد.

□ روش درست.

□ PCA در هنگام کاهش ابعاد، اطلاعات مربوط به خروجی را استفاده نمی‌کند.

□ برای مقابله با بیش‌برازش از تنظیم استفاده کنید.

$$\min_{\theta} \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2$$

□ طراحی یک سیستم یادگیری ماشین.

□ ایجاد مجموعه‌ی آموزشی به صورت $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$

□ استفاده از PCA به منظور کاهش ابعاد $x^{(i)}$ ها و به دست آوردن $z^{(i)}$ ها

□ اجرای مرحله‌ی آموزش بر روی $\{(z^{(1)}, y^{(1)}), (z^{(2)}, y^{(2)}), \dots, (z^{(m)}, y^{(m)})\}$

□ آزمایش فرضیه با استفاده از مجموعه‌ی آزمایشی: نگاشت $x_{test}^{(i)}$ به $z_{test}^{(i)}$ و محاسبه‌ی $h_{\theta}(z)$ برای هر یک از داده‌های مجموعه آزمایشی.

□ یک پرسش مهم. اگر فرآیند فوق را بدون استفاده از PCA انجام دهیم چه می‌شود؟

□ همواره ابتدا فرآیند بالا را بدون استفاده از PCA انجام دهید.

□ اگر به پاسخ مطلوب نرسیدید، آنگاه استفاده از PCA را آزمایش کنید.

کاربردها: فشردہ سازی

مجموعه‌ی آموزشی

۳۶



□ مجموعه‌ی آموزشی.

□ ۵۰۰۰ تصویر چهره (خاکستری)

□ ابعاد تصاویر.

□ ۳۲ در ۳۲ پیکسل

□ تعداد ویژگی‌ها.

□ ۱۰۲۴ ویژگی

اجرای PCA با $K = 100$

۳۷



□ ۳۶ مؤلفه‌ی اول.

تصاویر بازسازی شده

۳۸



تصاویر اولیه

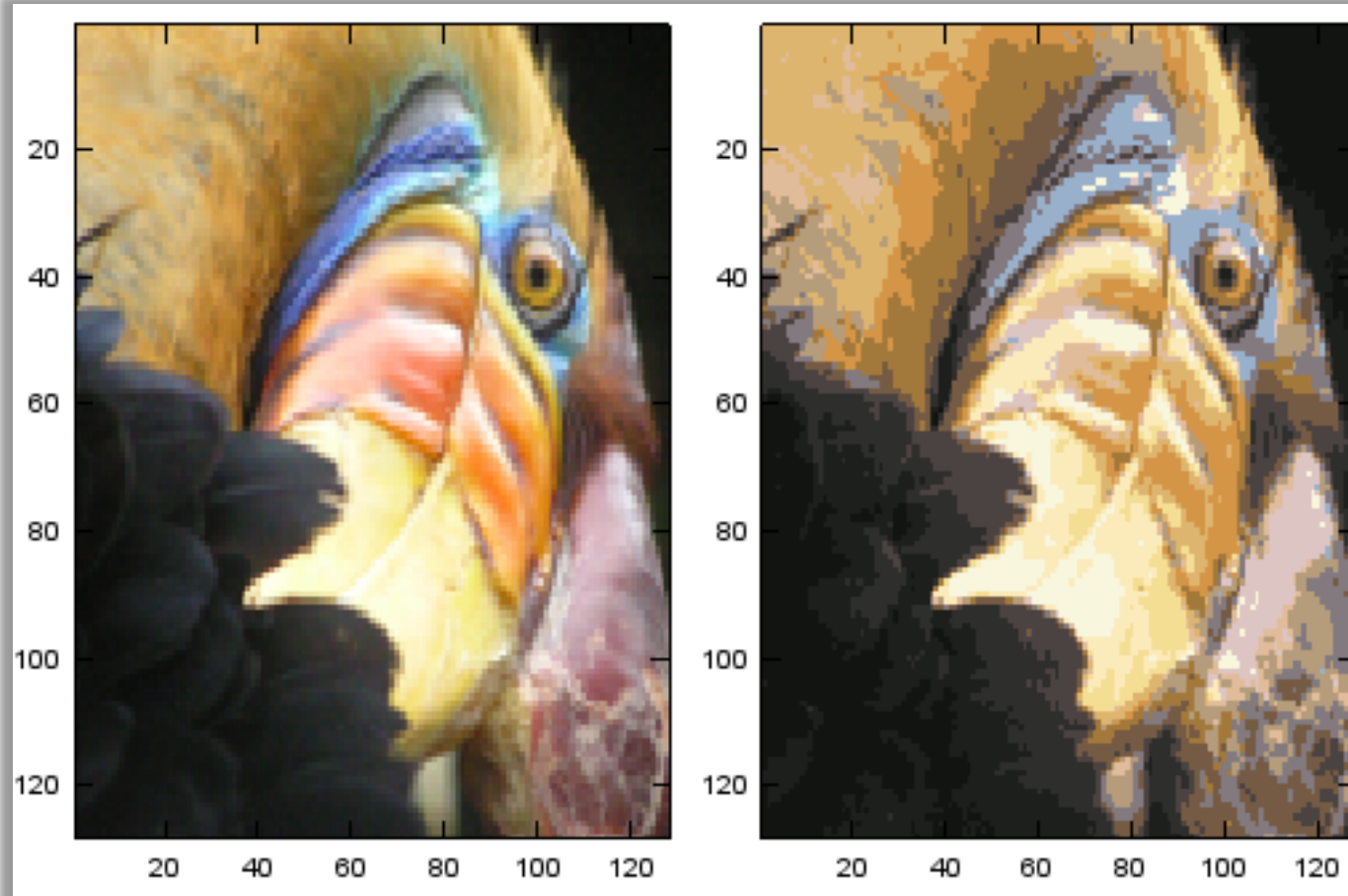


تصاویر بازسازی شده

کاربردها: تصویر نمودن داده‌ها

فشرده‌سازی با استفاده از k-means

۴۰

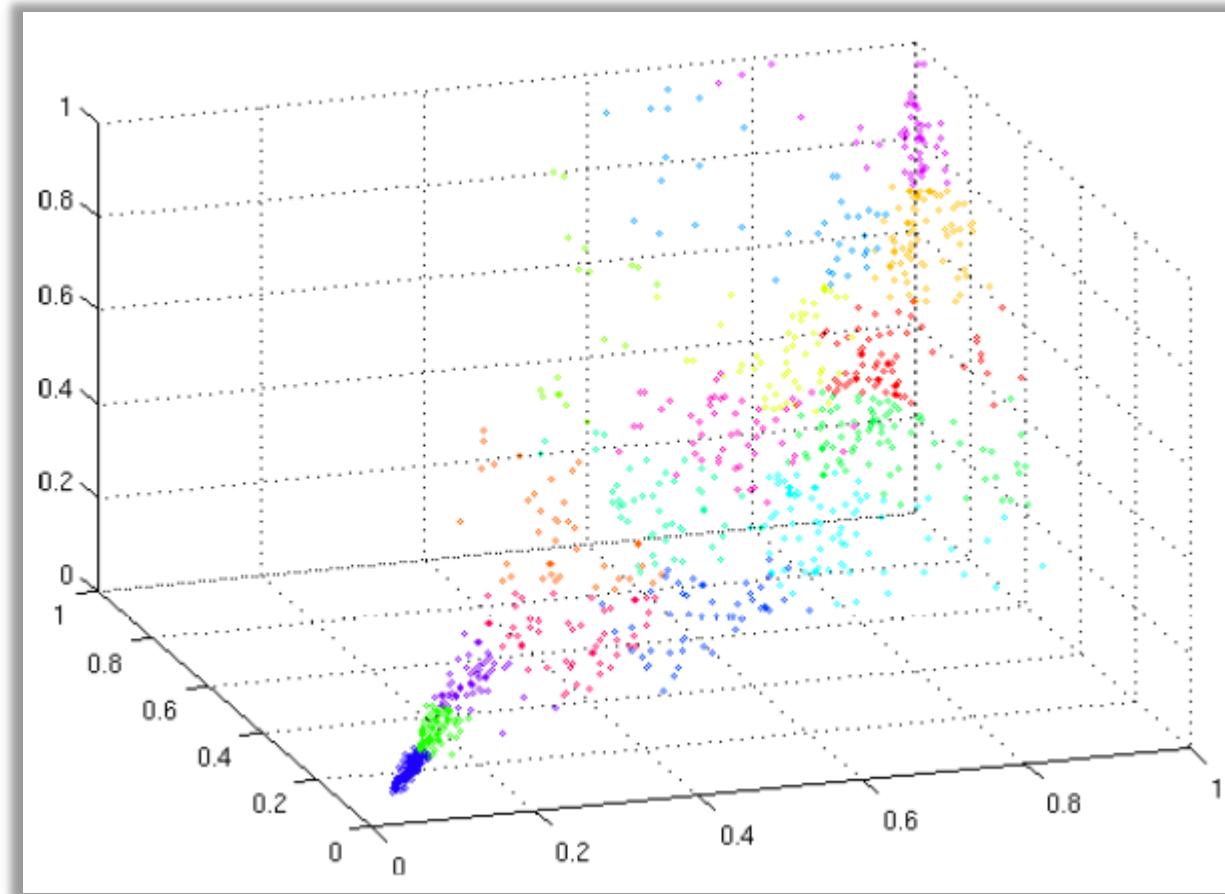


تصویر اولیه شامل هزاران رنگ (۱۲۸ در ۱۲۸)

تصویر فشرده شده شامل فقط ۱۶ رنگ

خوشه‌بندی پیکسل‌ها

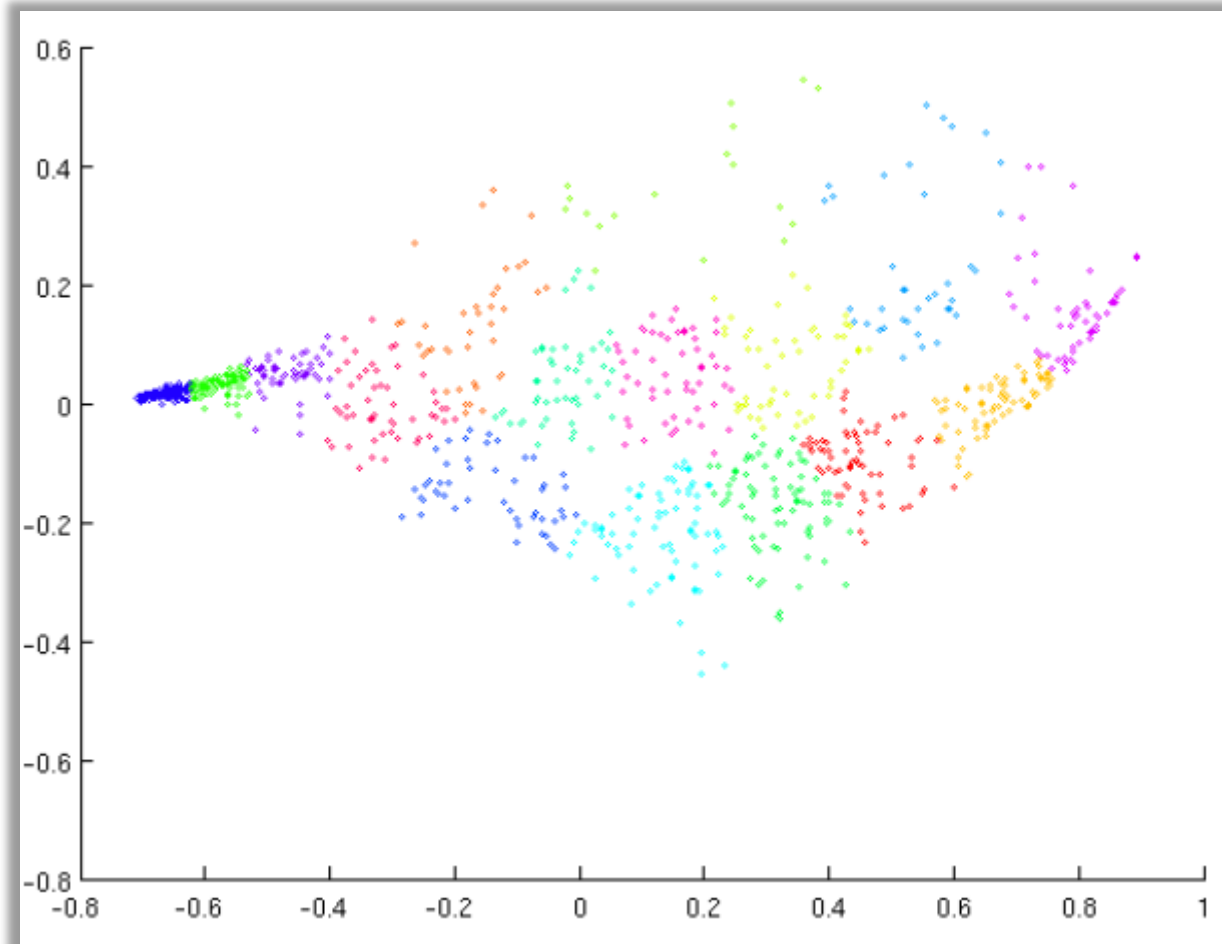
۴۱



خوشه‌بندی پیکسل‌ها به وسیله الگوریتم K-means

همان داده‌ها در فضای دو بعدی

۴۲



درک و تفسیر داده‌ها در فضای دو بعدی ساده‌تر است

تجزیه‌ی مقادیر منفرد

تجزیه‌ی مقادیر منفرد

۴۴

□ انگیزه.

- ساده‌سازی داده‌ها
- حذف نویز و افزودن
- بهبود نتایج الگوریتم

□ کاربردهای مثالی.

- جستجو و بازیابی اطلاعات [شاخص گذاری معنایی نهان]
- سیستم‌های توصیه‌گر

تجزیه‌ی مقادیر منفرد

۴۵

□ تجزیه‌ی مقادیر منفرد.

$$Data_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T$$

ماتریس مقادیر منفرد

□ ماتریس مقادیر منفرد.

- یک ماتریس قطری که در آن مقادیر منفرد به صورت کاهشی مرتب هستند.
- مقادیر منفرد از یک اندیس مانند r به بعد دارای مقدار صفر هستند.
- مقادیر منفرد ریشه‌ی دوم مقادیر ویژه‌ی ماتریس $Data \times Data^T$ هستند.

تجزیه‌ی مقادیر منفرد: مثال

۴۶

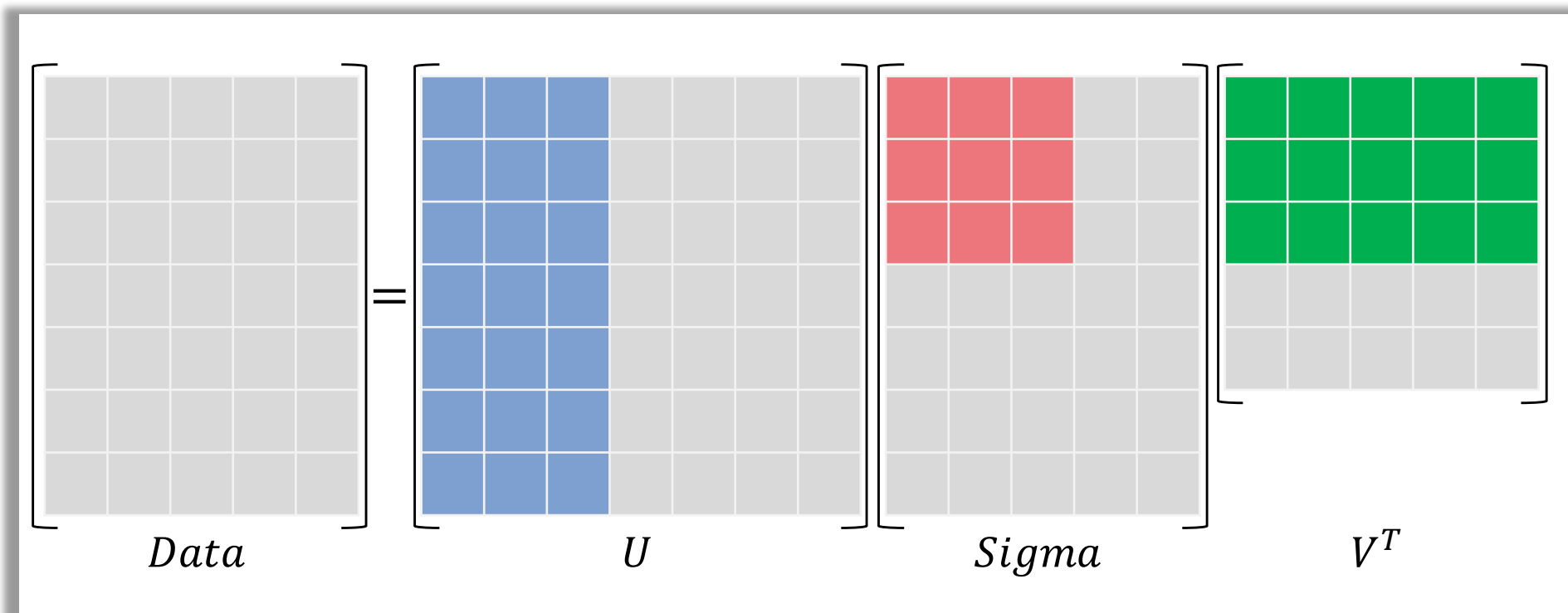
```
>> Data = [1 1 1 0 0; 2 2 2 0 0; 1 1 1 0 0; 5 5 5 0 0;  
           1 1 0 2 2; 0 0 0 3 3; 0 0 0 1 1];  
  
>> [U, Sigma, VT] = svd(Data);  
  
>> Sigma  
  
Sigma =  
Diagonal Matrix  
9.7214e+000      0      0      0      0  
      0 5.2940e+000      0      0      0  
      0      0 6.8423e-001      0      0  
      0      0      0 9.0541e-016      0  
      0      0      0      0 9.3837e-032  
      0      0      0      0      0  
      0      0      0      0      0  
  
>>
```

$$Data_{m \times n} \approx U_{m \times 3} \Sigma_{3 \times 3} V_{3 \times n}$$

تجزیه‌ی مقادیر منفرد: مثال

۴۷

$$Data_{m \times n} \approx U_{m \times 3} \Sigma_{3 \times 3} V_{3 \times n}^T$$



تجزیه‌ی مقادیر منفرد: مثال ($k = 1$)

۴۸

```
>> Sigma1x1 = Sigma(1:1, 1:1)
```

```
Sigma1x1 = 9.72140
```

```
>> Data_approx = U(:, 1:1) * Sigma1x1 * VT'(1:1, :)
```

```
1.05    1.05    0.98    0.06    0.06
2.01    2.01    1.96    0.12    0.12
1.00    1.00    0.98    0.06    0.06
5.03    5.03    4.90    0.30    0.30
0.76    0.76    0.73    0.04    0.04
0.12    0.12    0.12    0.00    0.00
0.04    0.04   -0.09    0.00    0.00
```

داره‌های اولیه

1	1	1	0	0
2	2	2	0	0
1	1	1	0	0
5	5	5	0	0
1	1	0	2	2
0	0	0	3	3
0	0	0	1	1

$SSE \approx 28.494$

تجزیه‌ی مقادیر منفرد: مثال ($k = 2$)

۴۹

```
>> Sigma2x2 = Sigma(1:2, 1:2)
```

```
Sigma2x2 =
```

```
Diagonal Matrix
```

```
    9.72140      0
      0    5.29398
```

```
>> Data_approx = U(:, 1:2) * Sigma2x2 * VT'(1:2, :)
```

```
1.00    1.00    0.99   -0.00   -0.00
2.00    2.00    1.98   -0.00   -0.00
1.00    1.00    0.99   -0.00   -0.00
5.02    5.02    4.94   -0.00   -0.00
0.77    0.77    0.00    2.03    2.03
0.14    0.14   -0.28    2.98    2.98
0.05    0.05   -0.09    0.99    0.99
```

داره‌های اولیه

1	1	1	0	0
2	2	2	0	0
1	1	1	0	0
5	5	5	0	0
1	1	0	2	2
0	0	0	3	3
0	0	0	1	1

$SSE \approx 0.46817$

تجزیه‌ی مقادیر منفرد: مثال ($k = 3$)

۵۰

```
>> Sigma3x3 = Sigma(1:3, 1:3)
```

```
Sigma3x3 =
```

```
Diagonal Matrix
```

```
    9.72140         0         0
         0    5.29398         0
         0         0    0.68423
```

```
>> Data_approx = U(:, 1:3) * Sigma3x3 * VT'(1:3, :)
```

```
1.    1.    1.    0.    0.
2.    2.    2.    0.    0.
1.    1.    1.   -0.   -0.
5.    5.    5.    0.    0.
1.    1.    0.    2.    2.
-0.    0.   -0.    3.    3.
-0.    0.   -0.    1.    1.
```

داره‌های اولیه

1	1	1	0	0
2	2	2	0	0
1	1	1	0	0
5	5	5	0	0
1	1	0	2	2
0	0	0	3	3
0	0	0	1	1

$$SSE \approx 1.2151e - 029$$

تعیین تعداد مقادیر منفرد

۵۱

□ تعیین یک تعداد مناسب برای مقادیر منفرد.

□ مشابه با تعیین تعداد مؤلفه‌های اصلی

□ یک روش تجربی. انتخاب کوچک‌ترین k به طوری که:

$$\frac{\sum_{i=1}^k s_{ii}^2}{\sum_{i=1}^n s_{ii}^2} \geq 0.90$$

$k = 1$. energy = 0.768

$k = 2$. energy = 0.996

$k = 3$. energy = 1.000

9.72	0	0	0	0
0	5.29	0	0	0
0	0	0.68	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

سیستم‌های توصیه‌گر

۵۲

□ سیستم‌های توصیه‌گر.

- «آمازون» بر اساس خریدهای گذشته‌ی مشتریان به آن‌ها اجناس خاصی را توصیه می‌کند.
- «نت فلیکس» بر اساس فیلم‌های تماشا شده، به مشتریان فیلم‌های خاصی را توصیه می‌کند.
- وبسایت‌های خبری به کاربران خود خواندن اخبار خاصی را پیشنهاد می‌کنند.

□ رویکرد استفاده شده. پالایش گروهی

- مقایسه‌ی داده‌های مربوط به یک کاربر با داده‌های مربوط به کاربران دیگر
- یافتن کاربران با سلیقه‌های مشابه و توصیه دادن به آن‌ها بر مبنای این مشابهت

شباهت مبتنی بر اقلام یا مبتنی بر کاربر؟

۵۳

□ مبتنی بر اقلام.

□ مثال: محاسبه‌ی شباهت دو فیلم مختلف بر اساس امتیازات داده شده از سوی کاربران

□ مبتنی بر کاربر.

□ مثال: محاسبه‌ی شباهت میان کاربران بر اساس امتیازات داده شده به وسیله‌ی آنها به فیلم‌های مختلف

□ کدام یک؟ محاسبه‌ی شباهت مبتنی بر اقلام یا مبتنی بر کاربر؟

□ بستگی به تعداد اقلام و تعداد کاربران دارد.

□ اگر تعداد کاربران بسیار بیشتر از تعداد اقلام است، محاسبه‌ی شباهت مبتنی بر اقلام بهتر است.

پيوست: مقادير ويژه و بردارهاي ويژه

مقادیر ویژه و بردارهای ویژه

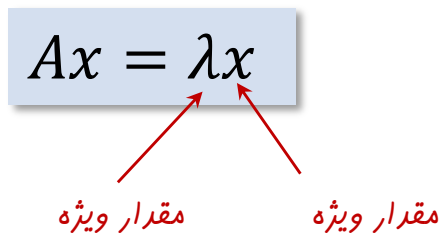
۵۵

□ حاصل ضرب ماتریس مربعی A را در بردار x در نظر بگیرید:

$$Ax = y$$

□ ماتریس A همانند یک تابع بردار x را به بردار جدید y تبدیل می‌کند.

□ بردار ویژه. بردار x یک بردار ویژه است اگر با بردار Ax موازی باشد:


$$Ax = \lambda x$$

مقدار ویژه مقدار ویژه

□ مقدار ویژه. در رابطه‌ی بالا، λ مقدار ویژه‌ی متناظر با بردار ویژه‌ی x است.

مقادیر ویژه و بردارهای ویژه

۵۶

□ مثال. اگر A یک ماتریس جایگشت به صورت زیر باشد:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

□ در این صورت:

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 \times \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \lambda = 1$$

$$A \begin{bmatrix} -1 \\ 1 \end{bmatrix} = (-1) \times \begin{bmatrix} -1 \\ 1 \end{bmatrix} \Rightarrow \lambda = -1$$

متعامد

$$\text{trace}(A) = 0 + 0 = 1 + (-1)$$

□ توجه. مجموع مقادیر ویژه با مجموع عناصر قطر اصلی (اثر) برابر است.

محاسبه‌ی مقادیر ویژه و بردارهای ویژه

۵۷

□ محاسبه‌ی مقادیر ویژه.

$$Ax = \lambda x \Rightarrow (A - \lambda I)x = 0$$

□ بنابراین ماتریس $A - \lambda x$ یک ماتریس منفرد است. [زیرا بردار پوچ دارد]

$$\det(A - \lambda I) = 0$$

□ در نتیجه:

□ مثال. محاسبه‌ی مقادیر ویژه

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix}\right) = \lambda^2 - \overset{\text{trace}(A)}{6}\lambda + \overset{\det(A)}{8} = 0 \Rightarrow \lambda = 4, 2$$

محاسبه‌ی مقادیر ویژه و بردارهای ویژه

۵۸

□ محاسبه‌ی بردارهای ویژه.

□ مثال.

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \Rightarrow \lambda_1 = 4, \lambda_2 = 2$$

$$(A - 4I)x_1 = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} x_1 = 0 \Rightarrow x_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$(A - 2I)x_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} x_2 = 0 \Rightarrow x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

□ مشاهده.

$$Ax = \lambda x \Rightarrow (A + 3I)x = Ax + 3x = \lambda x + 3x = (\lambda + 3)x$$

تجزیه‌ی ماتریس A: قطری سازی

۵۹

□ فرض کنید S ماتریسی باشد که ستون‌های آن بردارهای ویژه ماتریس A هستند.

$$\begin{aligned} AS &= A \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \end{bmatrix} \\ &= \begin{bmatrix} | & | & \cdots & | \\ \lambda_1 x_1 & \lambda_2 x_2 & \cdots & \lambda_n x_n \\ | & | & \cdots & | \end{bmatrix} \\ &= \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\ &= S\Lambda \end{aligned}$$

$$AS = S\Lambda \Rightarrow S^{-1}AS = \Lambda$$

$$AS = S\Lambda \Rightarrow A = S\Lambda S^{-1}$$

تجزیه ماتریس A: قطری سازی

۶۰

□ مشاهده. اگر ماتریس A را به توان دو برسانیم، مقادیر ویژه به توان دو می‌رسند و بردارهای ویژه تغییر نمی‌کنند.

$$A = S\Lambda S^{-1} \Rightarrow A^2 = S\Lambda S^{-1} S\Lambda S^{-1} = S\Lambda^2 S^{-1}$$

□ به طور کلی.

$$A = S\Lambda S^{-1} \Rightarrow A^k = S\Lambda^k S^{-1}$$

تجزیه‌ی ماتریس متقارن

۶۱

□ در یک ماتریس متقارن حقیقی:

□ مقادیر ویژه حقیقی هستند.

□ بردارهای ویژه متعامد نرمال هستند.

$$A = Q\Lambda Q^{-1} = Q\Lambda Q^T$$

$$A = \begin{bmatrix} | & | & \cdots & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & \cdots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} - & q_1^T & - \\ - & q_2^T & - \\ \vdots & \vdots & \vdots \\ - & q_n^T & - \end{bmatrix}$$
$$= \lambda_1 q_1 q_1^T + \lambda_2 q_2 q_2^T + \cdots + \lambda_n q_n q_n^T$$

□ مشاهده. هر ماتریس متقارن ترکیب خطی یک مجموعه از ماتریس‌های پروجکشن متعامد است.

ماتریس‌های مثبت معین متقارن

۶۲

□ ماتریس مثبت معین. ماتریس A مثبت معین است اگر به ازای هر بردار غیر صفر مانند x :

$$x^T A x > 0$$

□ در یک ماتریس مثبت معین متقارن:

□ مقادیر ویژه همگی مثبت هستند.

□ محورها همگی مثبت هستند.

□ دترمینان‌ها همگی (دترمینان زیرماتریس‌های مقدم) مثبت هستند.

□ مثال.

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 3 \end{bmatrix} \Rightarrow \lambda = 4 \pm \sqrt{5}, p_1 = 5, p_2 = \frac{11}{5}, \det(A) = 11$$