7. Largest Eigenvalue and Principal Directions

PCA of Covariance Matrix: The Largest Eigenvalues and the Principal Directions

(Caption will be displayed when you start playing the video.)

0:00 / 0:00 1.50x

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Empirical Variance in the Direction of the Top Eigenvector

1/1 point (ungraded)
Let $\mathbf{X}_1, \ldots, \mathbf{X}_n \in \mathbb{R}^d$ denote a data set. Let S denote the empirical covariance for this data set, and apply the decomposition theorem to write

$$S = PDP^T$$

where D is a diagonal matrix and $PP^T = I_d$. Further let's suppose that

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{pmatrix}, \quad P = \begin{pmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_d \\ \downarrow & \downarrow & \cdots & \downarrow \end{pmatrix}$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d \geq 0$ and $\mathbf{v}_1, \ldots, \mathbf{v}_d \in \mathbb{R}^d$.

What is

$$\mathbf{v}_1^T S \mathbf{v}_1?$$

- (selected) λ_1
- λ_d
- λ_1^2
- None of the above.

✔

Solution:

Observe that

$$
\begin{aligned}
\mathbf{v}_1^T S \mathbf{v}_1 &= \mathbf{v}_1^T P D P^T \mathbf{v}_1 \\
&= (1, 0, \ldots, 0) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
&= \lambda_1
\end{aligned}
$$

where we used that $PP^T = P^T P = I_d$.

Remark 1: The direction $\mathbf{v}_1$ (recall that $\|\mathbf{v}_1\|^2 = 1$) is the direction $\mathbf{w}$ that maximizes the empirical variance of the (one-dimensional) projected data set

$$\mathbf{w}^T \mathbf{X}_1, \mathbf{w}^T \mathbf{X}_2, \ldots, \mathbf{w}^T \mathbf{X}_n.$$

Remark 2: Similarly, we also have for $1 \le i \le d$

$$\mathbf{v}_i^T S \mathbf{v}_i = \lambda_i.$$

Submit

You have used 2 of 3 attempts

Answers are displayed within the problem

Is the Direction of Largest Empirical Variance Unique?

1/1 point (ungraded)

Consider the statistical set-up of the previous problem. In particular, recall that S denotes the empirical covariance matrix of the data set $\mathbf{X}_1, \ldots, \mathbf{X}_n$ and that S has eigenvalues $\lambda_1, \ldots, \lambda_d$ and corresponding eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_d$.

Unlike the previous problem, let's assume that we have **strict** inequalities

$$\lambda_1 > \lambda_2 > \cdots > \lambda_{d-1} > \lambda_d > 0.$$

We showed in the previous problem that $\mathbf{v}_1^T S \mathbf{v}_1 = \lambda_1$.

Does there exist a unit vector $\mathbf{w} \neq \mathbf{v}_1, \mathbf{w} \neq -\mathbf{v}_1$ such that

$$\mathbf{w}^T S \mathbf{w} \geq \lambda_1?$$

(Refer to the slides.)

- ○ Yes
- ● No

✔

Solution:

The correct answer is **"No"**. First observe that if $\mathbf{w}$ is a unit vector, then $P^T\mathbf{w}$ is also a unit vector. This is because

[Math Processing Error]

$$= (P^T\mathbf{w})^T P^T\mathbf{w}$$
$$= \mathbf{w}^T P P^T \mathbf{w}$$
$$= \mathbf{w}^T \mathbf{w}$$

[Math Processing Error]

$$= 1.$$

To go from the second to third line, we used that $PP^T = I_d$ and associativity of matrix multiplication.

Next, note that by the given decomposition

$$\mathbf{w}^T S\mathbf{w} = \mathbf{w}^T PDP^T\mathbf{w} = (P^T\mathbf{w})^T D\,(P^T\mathbf{w})\,.$$

But as $\mathbf{w}$ ranges over all unit vectors, we know that $P^T\mathbf{w}$ also ranges over all unit vectors. So if there exists $\mathbf{w} \neq \mathbf{v}_1$ such that $\mathbf{w}^T S\mathbf{w} = \lambda_1$, there must exist $\mathbf{b} \neq P^T\mathbf{v}_1 = (1, 0, \ldots, 0)^T$ such that $\mathbf{b}^T D\mathbf{b} = \lambda_1$. Observe that by matrix multiplication,

$$\mathbf{b}^T D\mathbf{b} = \sum_{i=1} \lambda_i (\mathbf{b}^i)^2 \le \lambda_1 (\mathbf{b}^1)^2 + \lambda_2\,(1 - \mathbf{b}_1^2)\,.$$

We also used that $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_d \ge 0$ and $\mathbf{b}_j^2 \le 1 - \mathbf{b}_1^2$ for all $j \neq 1$. Suppose $\mathbf{b}_1^2 \neq 1$ (so that $\mathbf{b} \neq \mathbf{v}_1$ and $1 - \mathbf{b}_1^2 > 0$), then we have

$$\mathbf{b}^T D\mathbf{b} \le \lambda_1 (\mathbf{b}^1)^2 + \lambda_2\,(1 - \mathbf{b}_1^2) < \lambda_1 (\mathbf{b}^1)^2 + \lambda_1\,(1 - \mathbf{b}_1^2) = \lambda_1,$$

where we used the strict inequality $\lambda_1 > \lambda_2$. Therefore, the equality case is **only** possible if $\mathbf{b} = (1, 0, \ldots, 0)^T$. Hence, we must also have $\mathbf{w} = \mathbf{v}_1$ if equality holds.

Submit

You have used 1 of 1 attempt

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: (Optional) Unit 8 Principal component analysis:(Optional) Lecture 23: Principal Component Analysis / 7. Largest Eigenvalue and Principal Directions

Add a Post

Show all posts ▼ by recent activity ▼

There are no posts in this topic yet.

✖