

پاسخ پرسش ها

پرسش اول: ماتریس کوواریانس داده ها در واقع از رابطه زیر بر حسب  $\tilde{X}$  بدست می آید:

$$Cov = \frac{1}{N-1} (\tilde{X} \cdot \tilde{X}^T)$$

که در آن  $N$  نشان دهنده تعداد داده ها می باشد. ماتریس کوواریانس داده ها در واقع میزان همبستگی بین دو متغیر تصادفی را نشان می دهد. وقتی مقادیر ویژه بیشینه و بردار های ویژه نظیر آن ها را برای این ماتریس بدست می آوریم در واقع در تلاش هستیم تا بیشترین واریانس داده ها (بیشترین اطلاعات داده) را روی دستگاه مختصات دوران یافته ای نگاشت کنیم. در واقع بردار های ویژه نقش داده های ما را دارند که با اعمال نگاشت ماتریسی کوواریانس (اعمال ضرایب واریانس و کوواریانس روی تمامی متغیر های تصادفی) ما را به بیشینه ضریبی از همان بردار داده ببرد. به همین منظور و از آن جایی که این بردار ها مستقل خطی هستند به عنوان پایه ای برای فضای مقصد انتخاب می شوند تا تمامی بردار های دیگر از روی ترکیب خطی آن ها ساخته شود.

پرسش دوم: در پاسخ به این سوال باید گفت که با این روش ما تمامی اطلاعات حاصل از متغیرهای تصادفی دیگر را به ترتیب از اولین متغیر تا آخرین متغیر سر شکن می کنیم به گونه ای که متغیر اول مثلاً ۴۰ درصد داده، متغیر دوم ۲۰ درصد داده، ... را داشته باشد. برای همین وقتی معادله ماتریسی بردار ویژه کوواریانس را به شکل دستگاه تشکیل می دهیم. مشاهده می کنیم که جمع مضارب از متغیرهای تصادفی (که مضارب همان کوواریانس های آن هستند) برابر مضربی از یک متغیر تصادفی می شود. وقتی بیشینه آن را انتخاب می کنیم در واقع حجم وسیع تری از داده را روی آن سر شکن می کنیم و به همین ترتیب برای متغیرهای دیگر نیز این به کار می رود. در پاسخ به این پرسش که ۳ ماتریس دیگر چه هستند باید گفت که ماتریس  $U$  همان بردارهای ویژه ماتریس کوواریانس هستند و ماتریس سیگما همان مقادیر منفرد ماتریس کوواریانس را دارد و در آخر نیز ماتریس  $V$  نشان دهنده بردارهای ویژه سطری برای ماتریسی هست که هر داریه آن نشان دهنده مجموع مربعات هر متغیر تصادفی از مقدار میانگین آن برای یک مشاهده است.

پرسش سوم: برای تبدیل چنین داده های تصادفی از بعد داده شده به بعد  $A$  باید  $A$  مقادیر منفرد اول که به ترتیب از بزرگترین به کوچکترین هستند را از روی ماتریس سیگما و انتخاب کرد سپس آن ها را به توان ۲ رساند تا مقادیر ویژه شان بدست آید. بعد از آن بردار ویژه نظیر آن ها را از روی ماتریس  $U$  انتخاب کرد و در یک ماتریس قرار داد و ترانهاده مزدوج آن را در ماتریس اصلی داده ضرب داخلی کرد. بدین ترتیب ماتریسی بدست می آید که داده ها را با بعد تقلیل یافته به  $A$  نشان می دهد.