A Simple, Math-Free Explanation of Principal Components Analysis (PCA)

Imagine you have a three-band aerial image (red, green, blue: RGB). You can display all of the bands at once on your RGB monitor, so you can see the inter-band variability. Because the bands are spectrally distinct, it is unlikely that any of the bands are highly correlated, e.g. that the red band is substantially similar to the green band. So the image's feature space plots (band bi-plots) will all show a fair bit of variability (i.e. a wide cloud of points rather than a more linear formation).

Now imagine that instead of a three-band image you have an image from a sensor that provides 1,000 contiguous bands across the visible and near infrared portions of the electromagnetic spectrum. Because there are so many bands in each portion of the spectrum (e.g. green), and the bands are spectrally adjacent, it is likely that many of the bands are highly correlated with other bands. For example, adjacent green bands may not show much inter-band variability (showing more linear feature space plots). In this case, we may choose to use PCA to capture the inter-band variability that exists, while also eliminating redundant bands.

PCA does this by "condensing"* the inter-band variability in the image to create a new image with bands that are ranked according to the fraction of total inter-band variability they represent. For example, PC-1, the first principal component, will represent the portions of the original image having the most inter-band variability, PC-2 will represent the next most variability, all the way to PC-*n* (where *n* is the number bands in the original image), which will likely have little information content remaining. It will look "noisy" because the major sources of variability in the image will have been represented in previous PC bands. Thus, feature space plots of the new PC bands (especially the first few) will show wider variability than many of the plots from the original image.

To summarize, PCA does three things for us:

1. Reduces image dimensionality (i.e. reduces the number of bands)
2. Reduces redundancy in the image
3. Accentuates the inter-band variability from the original image

*This process is accomplished using linear algebra, or "matrix math," and the PC bands are ranked by their eigenvalues.