There is a reason for using $n-1$ instead of n. The above definition makes $\mathrm{var}(x)$ an unbiased estimator of the variance of the random variable being sampled. However, we don't need to worry about this. Curious readers will find an explanation of these peculiar definitions in Epstein [57] (Chapter 14, Section 14.5) or in any decent statistics book.

Given two vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, the *sample covariance* (for short, *covariance*) of x and y is given by

$$\mathrm{cov}(x, y) = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{n-1}.$$

Example 23.7. If we take $x = (1, 3, -1)$ and $y = (0, 2, -2)$, we know from Example 23.6 that $x - \overline{x} = (0, 2, -2)$ and $y - \overline{y} = (-1, 0, 1)$. Thus, $\mathrm{cov}(x, y) = \frac{0(-1)+2(0)+(-2)(1)}{2} = -1$.

The covariance of x and y measures how x and y vary from the mean with respect to each other. Obviously, $\mathrm{cov}(x, y) = \mathrm{cov}(y, x)$ and $\mathrm{cov}(x, x) = \mathrm{var}(x)$.

Note that

$$\mathrm{cov}(x, y) = \frac{(x - \overline{x})^\top (y - \overline{y})}{n-1}.$$

We say that x and y are *uncorrelated* iff $\mathrm{cov}(x, y) = 0$.

Finally, given an $n \times d$ matrix X of n points X_i, for PCA to be meaningful, it will be necessary to translate the origin to the *centroid* (or *center of gravity*) μ of the X_i's, defined by

$$\mu = \frac{1}{n}(X_1 + \cdots + X_n).$$

Observe that if $\mu = (\mu_1, \ldots, \mu_d)$, then μ_j is the mean of the vector C_j (the jth column of X).

We let $X - \mu$ denote the *matrix* whose ith row is the centered data point $X_i - \mu$ ($1 \leq i \leq n$). Then the *sample covariance matrix* (for short, *covariance matrix*) of X is the $d \times d$ symmetric matrix

$$\Sigma = \frac{1}{n-1}(X - \mu)^\top (X - \mu) = (\mathrm{cov}(C_i, C_j)).$$

Example 23.8. Let $X = \begin{pmatrix} 1 & 1 \\ 3 & 2 \\ -1 & 3 \end{pmatrix}$, the 3×2 matrix whose columns are the vector x and y of Example 23.6. Then

$$\mu = \frac{1}{3}\left[(1, 1) + (3, 2) + (-1, 3)\right] = (1, 2),$$