Covariance of two random variables

- Height and wake-up time are uncorrelated, but height and weight are correlated.
- Covariance

  $\mathrm{Cov}(X, Y) = 0$ for X = height, Y = wake-up times

  $\mathrm{Cov}(X, Y) > 0$ for X = height, Y = weight

  - Definition:

$$\mathrm{Cov}(X,Y) \equiv C_{XY} \equiv \mathrm{E}\big((X-\mu_x)(Y-\mu_y)\big)$$

Question: If $\mathrm{Cov}(X, Y) < 0$ for two random variables X, Y, what would a scatterplot of samples from X, Y look like?

Question: If we add arbitrary constants to the random variables X, Y, how does the covariance change?

Estimating covariance from samples

Again, we assume that we do not know the underlying probability distributions. But consider we sample n times and estimate:

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{bmatrix}$$

$$Cov(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_x)(y_i - m_y)$$ "sample covariance"

Questions:

What is Cov(X, X) ? Cov(X, X) = Var(X)

How are Cov(X, Y) and Cov(Y, X) related? Cov(X, Y) = Cov(Y, X)

Estimating covariance in Matlab

- Sample mean:

$$x = [x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_n] \qquad\qquad m_x \leftarrow \texttt{m\_x}$$

$$y = [y_1 \quad y_2 \quad y_3 \quad \cdots \quad y_n] \qquad\qquad m_y \leftarrow \texttt{m\_y}$$

- Covariance

$$\mathrm{Cov}(X,Y) = \frac{1}{n}[x_1 - m_x \quad x_2 - m_x \quad \cdots \quad x_n - m_x] \begin{bmatrix} y_1 - m_y \\ y_2 - m_y \\ \vdots \\ y_n - m_y \end{bmatrix}$$

Method 1: `>> v = (1/n)*(x-m_x)*(y-m_y)'`

Method 2:

```
>> w = x-m_x
>> z = y-m_y
>> v = (1/n)*w*z'
```

Coefficient of Correlation

The coefficient of correlation is defined as:

$$\rho_{xy} \equiv \frac{\mathrm{E}\big((X-\mu_x)(Y-\mu_y)\big)}{\sigma_x \sigma_y} = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$$

Properties:

- $-1 \le C_{xy} \le 1$
- if $C_{xy} = 0$: X and Y uncorrelated
- if C_{xy} bigger/smaller zero : X and Y are positively/negatively correlated
- Advantage: we can multiply X and Y with arbitrary factors and C_{xy} stays the same.

Correlation of two random variables

Definition:

$$\mathrm{Corr}(X,Y) = \mathrm{E}\left(XY\right)$$

If X and Y have zero mean, this is the same as the covariance. If in addition, X and Y have variance of one this is the same as the coefficient of correlation.

Correlation, Covariance, Corr.Coeff.

$$\mathrm{Corr}(X,Y) = \mathrm{E}\left(XY\right)$$

Correlation

$$\mathrm{Cov}(X,Y) = \mathrm{E}\left((X-\mu_x)(Y-\mu_y)\right)$$

Covariance

$$\mathrm{C}_{xy} = \frac{\mathrm{E}\left((X-\mu_x)(Y-\mu_y)\right)}{\sigma_x \sigma_y} = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$$

Coefficient of Corr.

Covariance Matrix

Covariance: $$\mathrm{Cov}(A,B) = \mathrm{E}\big((A-\mu_A)(B-\mu_B)\big)$$

Now consider random vector: $X = (x_1, x_2, \cdots, x_n)^T$

We can compute covariance between two components, say between x_2 and x_5:

$$\mathrm{Cov}(x_2, x_5) = c_{25} = \mathrm{E}\big((x_2-\mu_{x_2})(x_5-\mu_{x_5})\big)$$

Doing this for all combinations gives us the elements of the covariance matrix:

$$\Sigma_X \equiv \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & & \\ \vdots & & \ddots & \vdots \\ & & & \\ c_{1n} & & \cdots & c_{nn} \end{pmatrix}$$

Properties of Covariance Matrix

$$\Sigma_X \equiv \begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ c_{12} & c_{22} & & \\ \vdots & & \ddots & \vdots \\ & & & \\ c_{1n} & & \cdots & c_{nn} \end{pmatrix} \qquad \mathrm{Cov}(x_i, x_j) = c_{ij} = \mathrm{E}\left((x_i - \mu_{x_i})(x_j - \mu_{x_j})\right)$$

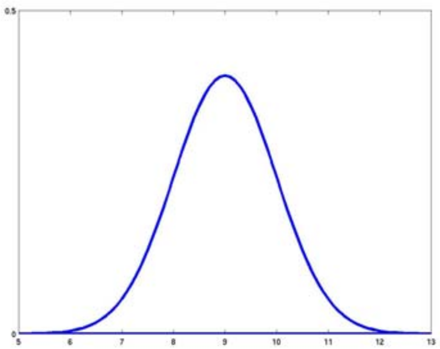
It is symmetric, because $c_{ij} = c_{ji}$.

Its diagonal elements are the individual variances: $c_{ii} = \sigma_i^{\;2} = \mathrm{var}(x_i)$.

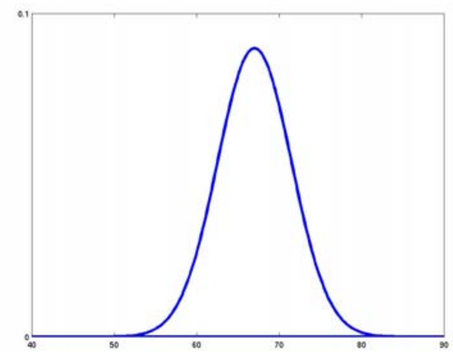
If it is *diagonal*, the x_i are all uncorrelated and we have:

$$\Sigma_X \equiv \begin{pmatrix} \sigma_1^{\;2} & 0 & \cdots & 0 \\ 0 & \sigma_2^{\;2} & & \\ \vdots & & \ddots & \vdots \\ & & & \\ 0 & & \cdots & \sigma_n^{\;2} \end{pmatrix}$$

Example covariance matrix



people's heights:
$X \sim \mathrm{N}(67, 20)$

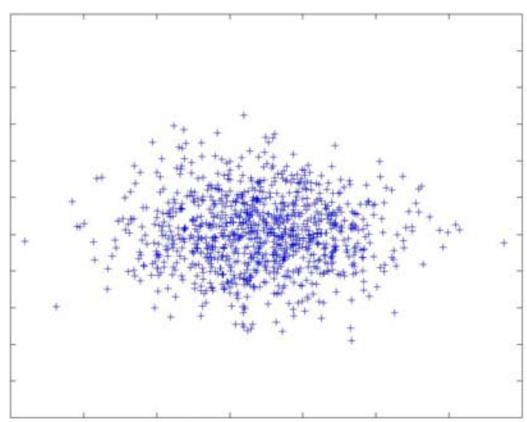


time people woke up this morning: $Y \sim \mathrm{N}(9, 1)$

Question: what is the covariance matrix of $V = (X\ Y)^{\mathrm{T}}$?

X and Y should be uncorrelated:

$$V = \begin{bmatrix} X \\ Y \end{bmatrix}$$



$$\begin{bmatrix} 20 & 0 \\ 0 & 1 \end{bmatrix}$$

Estimating the covariance matrix from samples (including Matlab code)

Sample n times and find mean of samples:

$$\begin{array}{c} \\ V = \end{array} \begin{array}{c} \begin{array}{cccc} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{array} \\ \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \end{bmatrix} \end{array} \qquad \mathbf{m} = \begin{bmatrix} m_x \\ m_y \end{bmatrix}$$

Find the covariance matrix:

$$Cov(V) = \frac{1}{n} \begin{bmatrix} x_1 - m_x & x_2 - m_x & \cdots & x_n - m_x \\ y_1 - m_y & y_2 - m_y & \cdots & y_n - m_y \end{bmatrix} \begin{bmatrix} x_1 - m_x & y_1 - m_y \\ x_2 - m_x & y_2 - m_y \\ \vdots & \vdots \\ x_n - m_x & y_n - m_y \end{bmatrix}$$

```
>> m = (1/n)*sum(v,2)
>> z = v - repmat(m,1,n)
>> v = (1/n)*z*z'
```

Gaussian distribution in *D* dimensions

Of course, the most important distribution can also be extended to higher dimensions. Recall that a 1-dimensional Gaussian is completely determined by its mean, μ, and its variance, σ^2:

$$X \sim \mathrm{N}(\mu, \sigma^2) \qquad p(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

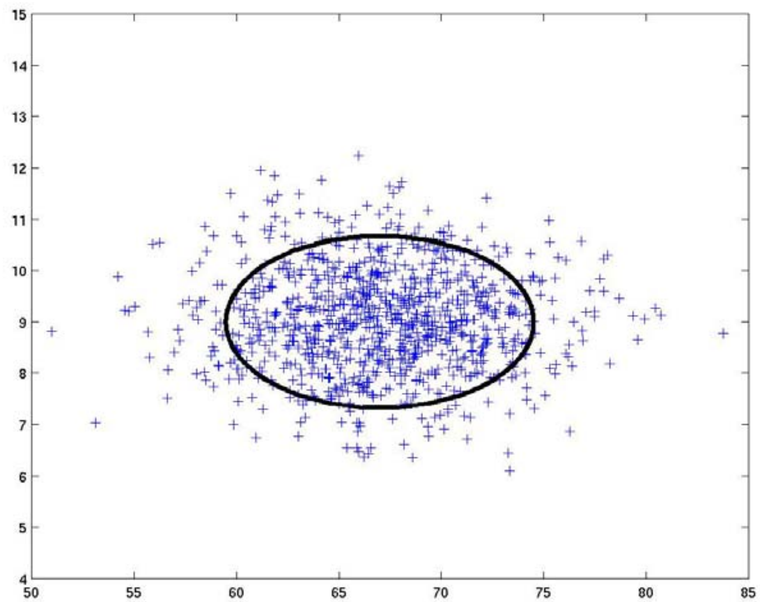
A D-dimensional Gaussian (multivariate Gaussian) is completely determined by its mean, μ, and its covariance matrix, Σ:

$$X \sim \mathrm{N}(\boldsymbol{\mu}, \Sigma) \qquad p(\mathbf{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}$$

Question: what happens when $D = 1$ for the D-dimensional Gaussian?

The Gaussian in *D* dimensions

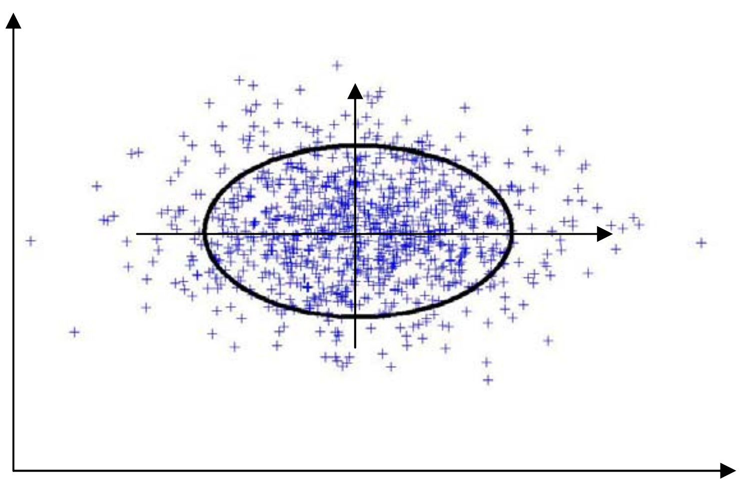
Question: What does a set of equiprobable points look like for a 2-dim. Gaussian? What for a D-dim. Gaussian?



In 2D, it's an ellipse. In *D* dimensions, it's an ellipsoid.

Equiprobable contours of a Gaussian

If a Gaussian random vector has covariance matrix that is diagonal (all of the variables are uncorrelated, then the axes of the ellipsoid are parallel to the coordinate axes.



Equiprobable contours of a Gaussian

If a Gaussian random vector has covariance matrix that is not diagonal (some of the variables are correlated), then the axes of the ellipsoid are perpendicular to each other, but are not parallel to the coordinate axes.