

Mean

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$$

Variance

$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}$$

$$V[X] = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}$$

Covariance

bivariate (x_i, y_i)

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}$$

$$\bar{y} = \sum_{i=1}^n \frac{y_i}{n}$$

$$\begin{aligned} C[x, y] &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y}) \\ &= \frac{1}{n} \left(\sum_{i=1}^n x_i y_i - \bar{y} \sum_{i=1}^n x_i - \bar{x} \sum_{i=1}^n y_i + \sum_{i=1}^n \bar{x} \bar{y} \right) \\ &= \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y} \end{aligned}$$

Covariance Matrix

$$\text{cov} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y})$$

$$\text{Cov}[x, y] = \begin{pmatrix} \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) (x_i - \bar{x}) & \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y}) \\ \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y}) & \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}) (y_i - \bar{y}) \end{pmatrix}$$

■ Example

```
x1 = {35, 5, 0, 20, 15};
x2 = {11, 12, 5, 9, 8};
x3 = {0.5, 0, 10, 4.5, 5};
{Mean[x1], Mean[x2], Mean[x3]}
{Variance[x1], Variance[x2], Variance[x3]}
```

```
{15, 9, 4.}
```

```
{ $\frac{375}{2}$ ,  $\frac{15}{2}$ , 16.375}
```

$$\bar{x}_1 = \sum_{i=1}^{\text{Length}[x_1]} \frac{x_1[[i]]}{\text{Length}[x_1]}$$

$$\sum_{i=1}^{\text{Length}[x_1]} \frac{(x_1[[i]] - \bar{x}_1)^2}{\text{Length}[x_1] - 1}$$

```
15
```

```
 $\frac{375}{2}$ 
```

$$\bar{x}_2 = \sum_{i=1}^{\text{Length}[x_2]} \frac{x_2[[i]]}{\text{Length}[x_2]}$$

$$\sum_{i=1}^{\text{Length}[x_2]} \frac{(x_2[[i]] - \bar{x}_2)^2}{\text{Length}[x_2] - 1}$$

```
9
```

```
 $\frac{15}{2}$ 
```

$$\left\{ \left\{ \sum_{i=1}^{\text{Length}[x_1]} \frac{(x_1[[i]] - \bar{x}_1)^2}{\text{Length}[x_1] - 1}, \sum_{i=1}^{\text{Length}[x_1]} \frac{(x_1[[i]] - \bar{x}_1)(x_2[[i]] - \bar{x}_2)}{\text{Length}[x_1] - 1} \right\}, \right.$$

$$\left. \left\{ \sum_{i=1}^{\text{Length}[x_1]} \frac{(x_1[[i]] - \bar{x}_1)(x_2[[i]] - \bar{x}_2)}{\text{Length}[x_1] - 1}, \sum_{i=1}^{\text{Length}[x_2]} \frac{(x_2[[i]] - \bar{x}_2)^2}{\text{Length}[x_2] - 1} \right\} \right\} // N // \text{MatrixForm}$$

```
( $\begin{pmatrix} 187.5 & 17.5 \\ 17.5 & 7.5 \end{pmatrix}$ )
```

```
{Part[{x1 -  $\bar{x}_1$ }.{x1 -  $\bar{x}_1$ }T, 1, 1], Part[{x1 -  $\bar{x}_1$ }.{x2 -  $\bar{x}_2$ }T, 1, 1]},
 {Part[{x1 -  $\bar{x}_1$ }.{x2 -  $\bar{x}_2$ }T, 1, 1], Part[{x2 -  $\bar{x}_2$ }.{x2 -  $\bar{x}_2$ }T, 1, 1]}} /
 (Length[x1] - 1) // N // MatrixForm
```

```
( $\begin{pmatrix} 187.5 & 17.5 \\ 17.5 & 7.5 \end{pmatrix}$ )
```

```

x1 = {1, 2, 3, 4, 5};
x2 = {3, 5, 2, 7, 1};
x3 = {1, 2, 3, 4, 5};
{Mean[x1], Mean[x2], Mean[x3]}
{Variance[x1], Variance[x2], Variance[x3]} // N

{3, 18/5, 3}

{2.5, 5.8, 2.5}

```

Correlation r

$$0 \leq |r_{xy}| \leq 1$$

$$\begin{aligned}
 r &= \frac{C[x, y]}{\sqrt{V[x] V[y]}} \\
 &= \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}} \sqrt{\sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n-1}}} \\
 &= \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}} \sqrt{\sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n-1}}} \\
 \rho_{xy} &= \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n-1}} \sqrt{\sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n-1}}}
 \end{aligned}$$

Cov[X,Y]

$$C_{xx} = \sigma_x^2$$

$$C_{yy} = \sigma_y^2$$

$$C_{xy} = \rho_{xy} \sigma_x \sigma_y$$

$$\text{Cov}[x, y] = \Sigma = E[\{X - \mu\} \cdot \{X - \mu\}^T] = \begin{pmatrix} C_{xx} & C_{xy} \\ C_{xy} & C_{yy} \end{pmatrix}$$

Probability

$$P[A] = \frac{n[A]}{n[S]}$$

Bayes

$$P[\omega_j | x] = \frac{P[x | \omega_j] \cdot P[\omega_j]}{P[x]}$$

$P[\omega_j]$: Prior

$P[\omega_j | \mathbf{x}]$: Posterior

$P[\mathbf{x} | \omega_j]$: Likelihood

$P[\mathbf{x}]$: Normalization

Expectation Probability

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$\bar{x} = \frac{w_1 x_1 + w_2 x_2 + \dots + w_n x_n}{w_1 + w_2 + \dots + w_n}$$

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} = \frac{1}{n} \sum_{\text{all } x} n_x x = \sum_{\text{all } x} x \frac{n_x}{n}$$

$$E[X] = \mu = \sum_{\text{all } x} x P[X] = \int_{-\infty}^{\infty} x f_x[x] dx$$

$$\frac{x_1}{\sum_{i=1}^{\text{Length}[x_1]} x_1[i]}$$

Probability Variance

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

$$= \sum_{\text{all } s} (x_i - \bar{x})^2 \frac{n_x}{n-1}$$

$$\sigma^2 = \sum_{\text{all } x} (x_i - \bar{x})^2 P[X]$$

$$\text{Var}[X] = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f_x[x] dx$$

$$\text{STD}[X] = \text{Var}[X]^{\frac{1}{2}}$$

Random Vector

Transpose $\begin{bmatrix} \text{ESC} \end{bmatrix} \text{tr} \begin{bmatrix} \text{ESC} \end{bmatrix}$

$$E[X] = [E[x_1] \ E[x_2] \ \dots \ E[x_n]]^T = [\mu_1 \ \mu_2 \ \dots \ \mu_n] = \mu$$

$$\text{Cov}[X] = \Sigma = E[(X - \mu)(X - \mu)^T]$$

$$\text{Cov}[X] =$$

$$\Sigma = E[(X - \mu)(X - \mu)^T] = E \begin{bmatrix} E[(x_1 - \mu)(x_1 - \mu)] & \dots & E[(x_1 - \mu)(x_N - \mu)] \\ \vdots & \ddots & \vdots \\ E[(x_N - \mu)(x_1 - \mu)] & \dots & E[(x_N - \mu)(x_N - \mu)] \end{bmatrix} = \begin{pmatrix} \sigma_1^2 & \dots & c_{1N} \\ \vdots & \ddots & \vdots \\ c_{1N} & \dots & \sigma_N^2 \end{pmatrix}$$

$$c_{ii} = \sigma_i^2$$

$$c_{ik} = \rho_{ik} \sigma_i \sigma_k$$

$$\Sigma = E[(X - \mu)(X - \mu)^T] = E[XX^T] - \mu\mu^T = S - \mu\mu^T$$

$$S = E[XX^T] = \begin{pmatrix} E[x_1 x_1] & \dots & E[x_1 x_N] \\ \vdots & \ddots & \vdots \\ E[x_N x_1] & \dots & E[x_N x_N] \end{pmatrix}$$

$$\Sigma = \Gamma R \Gamma^T = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \cdot \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & 1 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & 1 \end{pmatrix} \cdot \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}$$

Gaussian Distribution

$$N[\mu, \sigma] = P[X] = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$X = \{x_1, x_2, x_3\};$$

$$\mu = E[X];$$

$$\Sigma = E[(X - \mu)(X - \mu)^T];$$

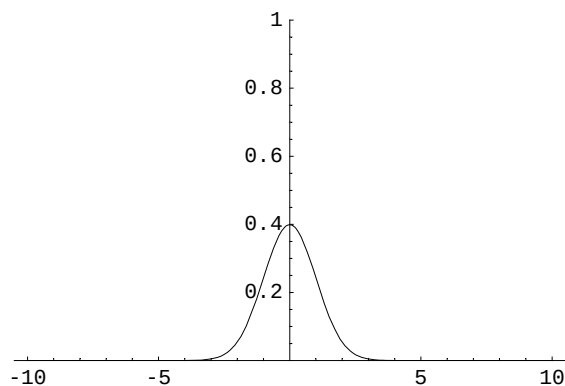
$$N[\mu, \Sigma] = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det[\Sigma]}} e^{-\frac{1}{2}(X-\mu)^T \Sigma^{-1} (X-\mu)}$$

■ Example

$$\mu_X = 0;$$

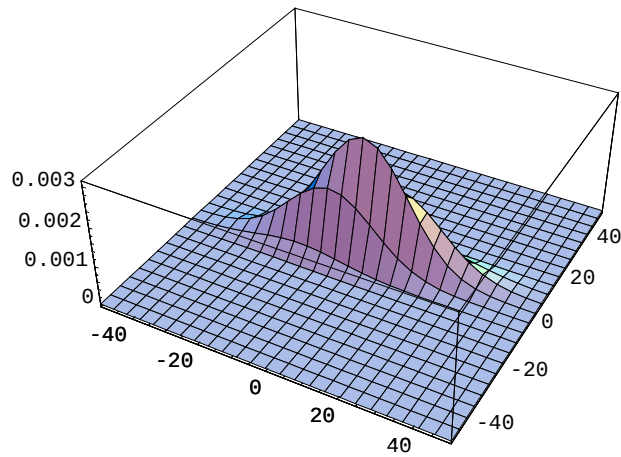
$$\sigma_X = 1;$$

$$\text{Plot}\left[\frac{1}{\sqrt{2\pi}\sigma_X} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}, \{x, -10, 10\}, \text{PlotRange} \rightarrow \{0, 1\}\right];$$



$\mu_1 = 0; \mu_2 = 0;$
 $\sigma_1 = 15; \sigma_2 = 4; \rho_{12} = 1/2;$

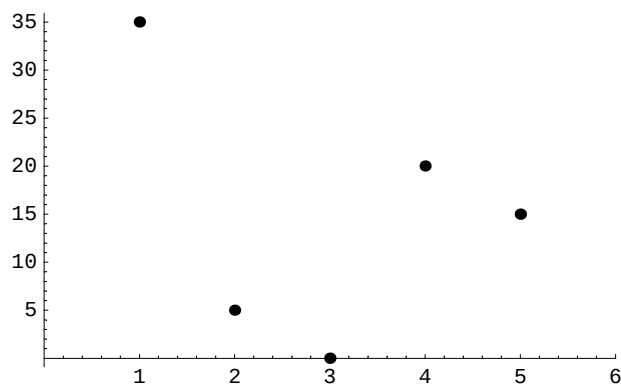
```
Plot3D[ $\frac{1}{2 \pi \sigma_1 \sigma_2 \sqrt{1 - \rho_{12}^2}} e^{-\frac{1}{2(1 - \rho_{12}^2)} \left( \left( \frac{x - \mu_1}{\sigma_1} \right)^2 - \frac{2 \rho_{12} (x - \mu_1) (y - \mu_2)}{\sigma_1 \sigma_2} + \left( \frac{y - \mu_2}{\sigma_2} \right)^2 \right)}$ ,  
{x, -50, 50}, {y, -50, 50}, PlotRange -> All];
```



Explain

 $x_1 = \{35, 5, 0, 20, 15\};$
 $n_1 = \text{Length}[x_1];$
 $P_1 = \{1, 1, 1, 1, 1\} / n_1;$

```
ListPlot[x1, PlotRange -> {{0, 6}, All}, PlotStyle -> PointSize[0.02]];
```


 $\text{Mean}[x_1]$
 $\text{Variance}[x_1]$
 15
 $\frac{375}{2}$

$$\blacksquare \mu = \sum_{\text{all } x} x P[x]$$

$$E_1 = \sum_{i=1}^{\text{Length}[x_1]} x_1[[i]] * P_1[[i]]$$

15

$$\blacksquare \sigma^2 = \sum_{\text{all } x} (x_i - \bar{x})^2 P[x]$$

$$\sigma_1 = \sum_{i=1}^{\text{Length}[x_1]} (x_1[[i]] - \bar{x}_1)^2 * P_1[[i]]$$

150

■ Total

$x_1 = \{35, 5, 0, 20, 15\};$

$x_2 = \{11, 12, 5, 9, 8\};$

$x_3 = \{1, 2, 3, 4, 5\};$

$n = \text{Length}[x_1];$

$P = \{1, 1, 1, 1, 1\} / (n);$

$$E_1 = \sum_{i=1}^n \text{Part}[x_1, i] * \text{Part}[P, i] // N$$

$$E_2 = \sum_{i=1}^n \text{Part}[x_2, i] * \text{Part}[P, i] // N$$

$$S_{11} = \sum_{i=1}^n (\text{Part}[x_1, i] - E_1)^2 * \text{Part}[P, i];$$

$$S_{12} = S_{21} = \sum_{i=1}^n (\text{Part}[x_1, i] - E_1) (\text{Part}[x_2, i] - E_2) * \text{Part}[P, i];$$

$$S_{22} = \sum_{i=1}^n (\text{Part}[x_2, i] - E_2)^2 * \text{Part}[P, i];$$

$$S_{11} = \text{Part}[\{x_1 - E_1\} \cdot \{x_1 - E_1\}^T, 1, 1] / (n - 1)$$

$$S_{12} = S_{21} = \text{Part}[\{x_1 - E_1\} \cdot \{x_2 - E_2\}^T, 1, 1] / (n - 1)$$

$$S_{22} = \text{Part}[\{x_2 - E_2\} \cdot \{x_2 - E_2\}^T, 1, 1] / (n - 1)$$

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} // N // \text{MatrixForm}$$

$$\rho_{12} = \frac{S_{12}}{\sqrt{S_{11}} \sqrt{S_{22}}} // N;$$

$$\text{Plot3D}\left[\frac{1}{2 \pi S_{11} S_{22} \sqrt{1 - \rho_{12}^2}} e^{-\frac{1}{2(1 - \rho_{12}^2)} \left(\left(\frac{x - E_1}{S_{11}} \right)^2 - \frac{2 \rho_{12} (x - E_1) (y - E_2)}{S_{11} S_{22}} + \left(\frac{y - E_2}{S_{22}} \right)^2 \right)}, \right. \\ \left. \{x, -500, 500\}, \{y, -500, 500\}, \text{PlotRange} \rightarrow \text{All}\right];$$

15.

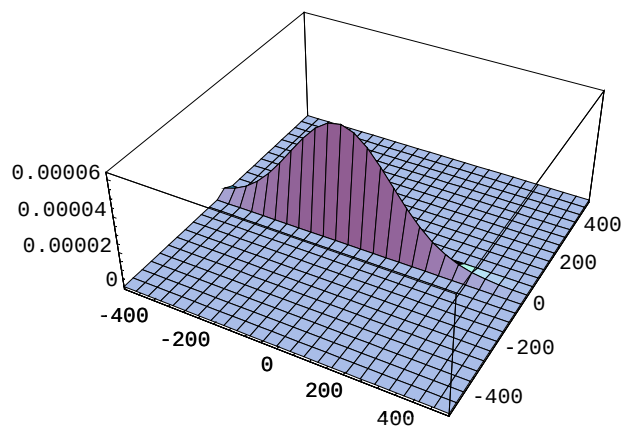
9.

187.5

17.5

7.5

$$\begin{pmatrix} 187.5 & 17.5 \\ 17.5 & 7.5 \end{pmatrix}$$



Multivariate Normal Distribution

k - dimensional

■ random vector $\mathbf{x} = (X_1, X_2, \dots, X_k)^T$

$$\mathbf{x} = [X_1, X_2, \dots, X_k]$$

$$\mathbf{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\mathbf{x} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

■ k-dimensional mean vector

$$\boldsymbol{\mu} = [E[X_1], E[X_2], \dots, E[X_k]]$$

■ k x k Covariance Matrix

$$\boldsymbol{\Sigma} = [\text{Cov}[X_i, X_j]]$$

■ Density

$$f_{\mathbf{x}}(x_1, \dots, x_k) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} e^{-\frac{1}{2} (\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}$$

| Σ | is determinant of Σ

■ Bivariate Case

$$k = \text{Rank}[\Sigma] = 2$$

$$\{X, Y\}^T$$

$$f[x, y] = \frac{1}{2 \pi \sigma_x \sigma_y \sqrt{1 - \rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x \sigma_y} \right)}$$

$$\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{pmatrix}$$

ρ is correlation

Derivation from Rotations

$$C_{xy} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}$$

$$\begin{pmatrix} u \\ v \end{pmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos[\theta] & -\sin[\theta] \\ \sin[\theta] & \cos[\theta] \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

$$C_{uv} = \begin{pmatrix} \sigma_{uu} & \sigma_{uv} \\ \sigma_{uv} & \sigma_{vv} \end{pmatrix} = T C_{xy} T^{-1}$$

$$C_{uv} = \begin{pmatrix} \cos[\theta] & -\sin[\theta] \\ \sin[\theta] & \cos[\theta] \end{pmatrix} \cdot \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{pmatrix} \cdot \begin{pmatrix} \cos[\theta] & -\sin[\theta] \\ \sin[\theta] & \cos[\theta] \end{pmatrix}^{-1} // \text{MatrixForm}$$

$$\begin{pmatrix} \sec[\theta] (\cos[\theta] \sigma_{xx} - \sin[\theta] \sigma_{xy}) + \csc[\theta] (\cos[\theta] \sigma_{xy} - \sin[\theta] \sigma_{yy}) & -\csc[\theta] (\cos[\theta] \sigma_{xx} - \sin[\theta] \sigma_{xy}) \\ \sec[\theta] (\sin[\theta] \sigma_{xx} + \cos[\theta] \sigma_{xy}) + \csc[\theta] (\sin[\theta] \sigma_{xy} + \cos[\theta] \sigma_{yy}) & -\csc[\theta] (\sin[\theta] \sigma_{xx} + \cos[\theta] \sigma_{xy}) \end{pmatrix}$$

$$\sigma_{uu} = \text{Part}[C_{uv}, 1, 1, 1];$$

$$\sigma_{vv} = \text{Part}[C_{uv}, 1, 2, 2];$$

$$\text{Solve}[\partial_\theta \sigma_{uu} == 0, \theta]$$

$$\text{Solve}[\partial_\theta \sigma_{vv} == 0, \theta]$$

Etc

Cal

$$\text{Inverse}\left[\begin{pmatrix} S_{1,1} & S_{1,2} \\ S_{2,1} & S_{2,2} \end{pmatrix}\right] = \begin{pmatrix} \frac{S_{2,2}}{-S_{1,2} S_{2,1} + S_{1,1} S_{2,2}} & -\frac{S_{1,2}}{-S_{1,2} S_{2,1} + S_{1,1} S_{2,2}} \\ -\frac{S_{2,1}}{-S_{1,2} S_{2,1} + S_{1,1} S_{2,2}} & \frac{S_{1,1}}{-S_{1,2} S_{2,1} + S_{1,1} S_{2,2}} \end{pmatrix}$$

$$\text{Det}[S] = \text{Det} \begin{bmatrix} S_{1,1} & S_{1,2} \\ S_{2,1} & S_{2,2} \end{bmatrix} = -S_{1,2} S_{2,1} + S_{1,1} S_{2,2}$$

$$\text{chol}[A] = R$$

$$R^T R = A$$

Covariance

■ 공분산 - 상관관계

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])]$$

$$\text{Cov}[X, Y] = E[XY] - E[X]E[Y]$$

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])^T]$$

$$\text{Cov}[X, Y] = E[XY^T] - E[X]E[Y]^T$$

$$\text{Cov}[X, 0] = 0$$

$$\text{Cov}[X, X] = \text{Var}[X]$$

$$\text{Cov}[X, Y] = \text{Cov}[Y, X]$$

$$\text{Cov}[aX, bY] = ab \text{Cov}[X, Y]$$

$$\text{Cov}[X + a, Y + b] = \text{Cov}[X, Y]$$

$$\text{Cov}[aX + bY, cW + dV] = ac \text{Cov}[X, W] + ad \text{Cov}[X, V] + bc \text{Cov}[Y, W] + bd \text{Cov}[Y, V]$$

$$\text{Cov}[X_i, X_j] = E[(X_i - E[X_i])(X_j - E[X_j])] = E[X_i X_j] - E[X_i] \cdot E[X_j]$$

■ 공분산행렬

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

$$\sum_{i,j} = \text{Cov}[X_i, X_j] = E[(X_i - E[X_i])(X_j - E[X_j])] = E[(X_i - \mu_i)(X_j - \mu_j)]$$

$$\sum = \begin{pmatrix} E[(X_1 - \mu_1)(X_1 - \mu_1)] & E[(X_1 - \mu_1)(X_2 - \mu_2)] & \cdots & E[(X_1 - \mu_1)(X_n - \mu_n)] \\ E[(X_2 - \mu_2)(X_1 - \mu_1)] & E[(X_2 - \mu_2)(X_2 - \mu_2)] & \cdots & E[(X_2 - \mu_2)(X_n - \mu_n)] \\ \vdots & \vdots & \ddots & \vdots \\ E[(X_n - \mu_n)(X_1 - \mu_1)] & E[(X_n - \mu_n)(X_2 - \mu_2)] & \cdots & E[(X_n - \mu_n)(X_n - \mu_n)] \end{pmatrix}$$

$$\sum = E[(X - E[X])(X - E[X])^T]$$

$$\sigma^2 = \text{Var}[X] = E[(X - E[X])^2] = E[(X - \mu)^2]$$

$$\mu = E[X]$$

$$\text{Var}[X] = \text{Cov}[X] = E[(X - E[X])(X - E[X])^T]$$

$$\text{Cov}[X, Y] = E[(X - \mu)(Y - \nu)^T]$$

$$\sum = E[(X - E[X])(X - E[X])^T], \mu = E[X]$$

$$\Sigma = E[XX^T] - \mu\mu^T$$

$$\text{Var}[AX + a] = A \text{Var}[X] A^T$$

$$\text{Cov}[X, Y] = \text{Cov}[Y, X]$$

$$\text{Cov}[X_1 + X_2, Y] = \text{Cov}[X_1, Y] + \text{Cov}[X_2, Y]$$

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Cov}[X, Y] + \text{Cov}[Y, X] + \text{Var}[Y]$$

$$\text{Cov}[AX, B^T Y] = A \text{Cov}[X, Y] B$$

X, X_1, X_2 : $p \times 1$ vector

Y : $q \times 1$ vector

a : $q \times 1$ vector

A, B : $q \times p$ matrices

$$c^T \Sigma = \text{Cov}[c^T X, X]$$

$$d^T \Sigma c = \text{Cov}[d^T X, c^T X]$$

$$\text{Var}[b^T X] = b^T \text{Var}[X] b$$

Matrix, Derivative, det, tr

$$\text{Inverse}\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right] = \begin{pmatrix} \frac{d}{-b c + a d} & -\frac{b}{-b c + a d} \\ -\frac{c}{-b c + a d} & \frac{a}{-b c + a d} \end{pmatrix}$$

$$\text{Det}\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right] = -b c + a d$$

$$\text{Tr}\left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right] = a + d$$

- The product of the eigenvalues is equal to the determinant of A

$$\det[A] = \prod_{i=1}^{N_\lambda} \lambda_i^{n_i}$$

- The sum of the eigenvalues is equal to the trace of A

$$\text{tr}[A] = \sum_{i=1}^{N_\lambda} n_i \lambda_i$$

■ Vector calculus

The tangent vector to a curve $x : \mathbb{R} \rightarrow \mathbb{R}^n$

$$\frac{\partial x}{\partial t} = \begin{pmatrix} \frac{\partial x_1}{\partial t} \\ \vdots \\ \frac{\partial x_n}{\partial t} \end{pmatrix}$$

The gradient of a scalar function $f : \mathbb{R}^n \rightarrow \mathbb{R}$

$$\frac{\partial f}{\partial x} = \left(\frac{\partial f}{\partial x_1} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right)$$

The directional derivative of f in the direction of v is then

$$\nabla_v f = \frac{\partial f}{\partial x} v$$

The pushforward of differential of a function $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is described by the Jacobian matrix

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix}$$

The pushforward along f of a vector v in $\mathbb{R}^m$

$$df(v) = \frac{\partial f}{\partial x} v$$

■ Matrix calculus

The tangent vector of a curve $F : \mathbb{R} \rightarrow M(n, m)$

$$\frac{\partial F}{\partial t} = \begin{pmatrix} \frac{\partial F_{1,1}}{\partial t} & \cdots & \frac{\partial F_{1,m}}{\partial t} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{n,1}}{\partial t} & \cdots & \frac{\partial F_{n,m}}{\partial t} \end{pmatrix}$$

The gradient of a scalar function $f : M(n, m) \rightarrow \mathbb{R}$

$$\frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f}{\partial x_{1,1}} & \cdots & \frac{\partial f}{\partial x_{1,m}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_{n,1}} & \cdots & \frac{\partial f}{\partial x_{n,m}} \end{pmatrix}$$

$$\nabla_Y f = \text{tr} \left[\frac{\partial f}{\partial x} Y \right]$$

$F : M(n, m) \rightarrow M(p, q)$

$$\frac{\partial F}{\partial x} = \begin{pmatrix} \frac{\partial F}{\partial x_{1,1}} & \cdots & \frac{\partial F}{\partial x_{1,m}} \\ \vdots & \ddots & \vdots \\ \frac{\partial F}{\partial x_{n,1}} & \cdots & \frac{\partial F}{\partial x_{n,m}} \end{pmatrix}$$

$$dF(Y) = \text{tr} \left(\frac{\partial F}{\partial x} Y \right)$$

$$\frac{\partial F}{\partial x} = \begin{pmatrix} \frac{\partial F_{1,1}}{\partial x} & \cdots & \frac{\partial F_{1,m}}{\partial x} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{n,1}}{\partial x} & \cdots & \frac{\partial F_{n,m}}{\partial x} \end{pmatrix}$$

$$\frac{\partial F}{\partial x} = \begin{pmatrix} \frac{\partial F}{\partial x_{1,1}} & \cdots & \frac{\partial F}{\partial x_{1,m}} \\ \vdots & \ddots & \vdots \\ \frac{\partial F}{\partial x_{n,1}} & \cdots & \frac{\partial F}{\partial x_{n,m}} \end{pmatrix}$$

Jacobian matrix

$$DF(x) = \frac{\partial \text{vec } F(x)}{\partial (\text{vec } x)^T}$$

■ Chain rule

$$\frac{\partial Z}{\partial X} = \frac{\partial Z}{\partial Y} \frac{\partial Y}{\partial X}$$

■ Product rule

$$\frac{\partial (YZ)}{\partial X} = \frac{\partial Y}{\partial X} Z + Y \frac{\partial Z}{\partial X}$$

■ Derivative of linear functions

$$\frac{\partial \mathbf{a}^T \mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}^T \mathbf{a}}{\partial \mathbf{x}} = \mathbf{a}^T$$

$$\frac{\partial \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}^T \mathbf{A}}{\partial \mathbf{x}} = \mathbf{A}$$

■ Derivative of quadratic functions

$$\frac{\partial \mathbf{x}^T \mathbf{A} \mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}^T (\mathbf{A}^T + \mathbf{A}) = \mathbf{x}^T \mathbf{A}^T + \mathbf{x}^T \mathbf{A}$$

$$\frac{\partial (\mathbf{A} \mathbf{x} + \mathbf{b})^T \mathbf{C} (\mathbf{D} \mathbf{x} + \mathbf{e})}{\partial \mathbf{x}} = (\mathbf{D} \mathbf{x} + \mathbf{e})^T \mathbf{C}^T \mathbf{A} + (\mathbf{A} \mathbf{x} + \mathbf{b})^T \mathbf{C} \mathbf{D}$$

$$\frac{\partial \|\mathbf{x} - \mathbf{a}\|}{\partial \mathbf{x}} = \frac{(\mathbf{x} - \mathbf{a})^T}{\|\mathbf{x} - \mathbf{a}\|}$$

■ Derivative of matrix traces

$$\frac{\partial \text{tr}(\mathbf{A} \mathbf{C})}{\partial \mathbf{A}} = \mathbf{C}^T$$

$$\frac{\partial \text{tr}(\mathbf{A} \mathbf{X} \mathbf{B})}{\partial \mathbf{X}} = \frac{\partial \text{tr}(\mathbf{B}^T \mathbf{X}^T \mathbf{A}^T)}{\partial \mathbf{X}} = \mathbf{B} \mathbf{A}$$

$$\frac{\partial \text{tr}(\mathbf{A} \mathbf{B} \mathbf{A}^T)}{\partial \mathbf{A}} = 2 \mathbf{A} \mathbf{B}$$

$$\frac{\partial \text{tr}(\mathbf{A} \mathbf{X} \mathbf{B} \mathbf{X}^T \mathbf{C})}{\partial \mathbf{X}} = \mathbf{B} \mathbf{X}^T \mathbf{C} \mathbf{A} + \mathbf{B}^T \mathbf{X}^T \mathbf{A}^T \mathbf{C}^T$$

■ Derivative of matrix determinant

$$\frac{\partial \det \mathbf{X}}{\partial \mathbf{X}} = \text{adj} \mathbf{X} = \det \mathbf{X} \cdot \mathbf{X}^{-1}$$

Jacobian matrix

$$\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\mathbf{J} = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \vdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

$$\mathbf{J}_F(x_1, \dots, x_n) \text{ and } \frac{\partial (y_1, \dots, y_m)}{\partial (x_1, \dots, x_n)}$$

$$\mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{p}) + \mathbf{J}_F(\mathbf{p})(\mathbf{x} - \mathbf{p}) + o(\|\mathbf{x} - \mathbf{p}\|)$$

■ ex1

$$x_1 = r \sin[\theta] \cos[\phi]$$

$$x_2 = r \sin[\theta] \sin[\phi]$$

$$x_3 = r \cos[\theta]$$

$$\mathbf{J}_F(r, \theta, \phi) = \begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \theta} & \frac{\partial x_1}{\partial \phi} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \theta} & \frac{\partial x_2}{\partial \phi} \\ \frac{\partial x_3}{\partial r} & \frac{\partial x_3}{\partial \theta} & \frac{\partial x_3}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \sin[\theta] \cos[\phi] & r \cos[\theta] \cos[\phi] & -r \sin[\theta] \sin[\phi] \\ \sin[\theta] \sin[\phi] & r \cos[\theta] \sin[\phi] & r \sin[\theta] \cos[\phi] \\ \cos[\theta] & -r \sin[\theta] & 0 \end{pmatrix}$$

■ ex2

$$x = r \cos[\phi]$$

$$y = r \sin[\phi]$$

$$\mathbf{J}_F(r, \phi) = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \phi} \end{pmatrix} = \begin{pmatrix} \cos[\phi] & -r \sin[\phi] \\ \sin[\phi] & r \cos[\phi] \end{pmatrix}$$

$$\iint_A dx dy = \iint_B r dr d\phi$$

■ ex3

$$\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$y_1 = F_1(x_1, \dots, x_n)$$

$$\vdots$$

$$y_m = F_m(x_1, \dots, x_n)$$

$$\mathbf{J}_F(x_1, \dots, x_n) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \vdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

■ cholesky factorization

$$\mathbf{R} = \text{chol}(\mathbf{A})'$$

$$\mathbf{R} * \mathbf{R}' = \mathbf{A}$$

$$\mathbf{R} = \text{chol}(\mathbf{A})$$

$$\mathbf{R}' * \mathbf{R} = \mathbf{A}$$