Random vectors and dependence modelling

Random vectors and dependence modelling

Goal: model the risk factor changes $X_n = (X_{n,1}, X_{n,2}, \ldots, X_{n,d})$
Assumption: $X_{n,i}$ and $X_{n,j}$ are dependent but $X_{n,i}$ und $X_{n\pm k,j}$ are independent fot $k \in \mathbb{N}$, $k \neq 0$, $1 \leq i,j \leq d$.

Random vectors and dependence modelling

Goal: model the risk factor changes $X_n = (X_{n,1}, X_{n,2}, \ldots, X_{n,d})$
Assumption: $X_{n,i}$ and $X_{n,j}$ are dependent but $X_{n,i}$ und $X_{n\pm k,j}$ are independent fot $k \in \mathbb{N}$, $k \neq 0$, $1 \leq i, j \leq d$.

Random vectors and dependence modelling

Goal: model the risk factor changes $X_n = (X_{n,1}, X_{n,2}, \ldots, X_{n,d})$
Assumption: $X_{n,i}$ and $X_{n,j}$ are dependent but $X_{n,i}$ und $X_{n\pm k,j}$ are independent fot $k \in \mathbb{N}$, $k \neq 0$, $1 \leq i, j \leq d$.

A d-dimensional random vector $X = (X_1, X_2, \ldots, X_d)^T$ is uniquely specified by its (multivariate) cumulative distribution function (c.d.f.) F:

$$F(x) : F(x_1, x_2, \ldots, x_d) := P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_d \leq x_d) = P(X \leq x).$$

Random vectors and dependence modelling

Goal: model the risk factor changes $X_n = (X_{n,1}, X_{n,2}, \ldots, X_{n,d})$
Assumption: $X_{n,i}$ and $X_{n,j}$ are dependent but $X_{n,i}$ und $X_{n\pm k,j}$ are independent fot $k \in \mathbb{N}$, $k \neq 0$, $1 \leq i, j \leq d$.

A d-dimensional random vector $X = (X_1, X_2, \ldots, X_d)^T$ is uniquely specified by its (multivariate) cumulative distribution function (c.d.f.) F:

$$F(x) : F(x_1, x_2, \ldots, x_d) := P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_d \leq x_d) = P(X \leq x).$$

The i-th marginal distribution F_i of F is the distribution function of X_i given as follows:

$$F_i(x_i) = P(X_i \leq x_i) = F(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty)$$

Random vectors and dependence modelling

Goal: model the risk factor changes $X_n = (X_{n,1}, X_{n,2}, \ldots, X_{n,d})$
Assumption: $X_{n,i}$ and $X_{n,j}$ are dependent but $X_{n,i}$ und $X_{n\pm k,j}$ are independent fot $k \in \mathbb{N}$, $k \neq 0$, $1 \leq i, j \leq d$.

A d-dimensional random vector $X = (X_1, X_2, \ldots, X_d)^T$ is uniquely specified by its (multivariate) cumulative distribution function (c.d.f.) F:

$$F(x) : F(x_1, x_2, \ldots, x_d) := P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_d \leq x_d) = P(X \leq x).$$

The i-th marginal distribution F_i of F is the distribution function of X_i given as follows:

$$F_i(x_i) = P(X_i \leq x_i) = F(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty)$$

The distribution function F is continuous if there exists a non-negative function $f \geq 0$, such that

$$F(x_1, x_2, \ldots, x_d) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \ldots \int_{-\infty}^{x_d} f(u_1, u_2, \ldots, u_d) du_1 du_2 \ldots du_d$$

f is then called the *(multivariate) density function* (d.f.) of F.

Ramdom vectors (contd.)

Ramdom vectors (contd.)

The components of X are independent iff $F(x) = \Pi_{i=1}^{d} F_i(x_i)$. If the d.f. f and the marginal d.f. f_i, $1 \leq i \leq d$, exist, then the components of X are independent iff

$$f(x) = \Pi_{i=1}^{d} f_i(x_i)$$

Ramdom vectors (contd.)

The components of X are independent iff $F(x) = \Pi_{i=1}^{d} F_i(x_i)$. If the d.f. f and the marginal d.f. f_i, $1 \le i \le d$, exist, then the components of X are independent iff

$$f(x) = \Pi_{i=1}^{d} f_i(x_i)$$

A random vector can be uniquely characterized in terms of its characteristic function $\phi_X(t)$:

$$\phi_X(t) := E(\exp\{it^T X\}), t \in \mathbb{R}^d$$

Ramdom vectors (contd.)

The components of X are independent iff $F(x) = \Pi_{i=1}^{d} F_i(x_i)$. If the d.f. f and the marginal d.f. f_i, $1 \leq i \leq d$, exist, then the components of X are independent iff

$$f(x) = \Pi_{i=1}^{d} f_i(x_i)$$

A random vector can be uniquely characterized in terms of its characteristic function $\phi_X(t)$:

$$\phi_X(t) := E(\exp\{it^T X\}), t \in \mathbb{R}^d$$

If $E(X_k^2) < \infty$ for all k, the the covariance (matrix) of X exists and is given es

$$Cov(X) = E((X - E(X))(X - E(X))^T)$$

Ramdom vectors (contd.)

The components of X are independent iff $F(x) = \Pi_{i=1}^{d} F_i(x_i)$. If the d.f. f and the marginal d.f. f_i, $1 \leq i \leq d$, exist, then the components of X are independent iff

$$f(x) = \Pi_{i=1}^{d} f_i(x_i)$$

A random vector can be uniquely characterized in terms of its characteristic function $\phi_X(t)$:

$$\phi_X(t) := E(\exp\{it^T X\}), t \in \mathbb{R}^d$$

If $E(X_k^2) < \infty$ for all k, the the covariance (matrix) of X exists and is given es

$$Cov(X) = E((X - E(X))(X - E(X))^T)$$

For an n-dimensional random vector X, a constant matrix $B \in \mathbb{R}^{n \times n}$ and a constant vector $b \in \mathbb{R}^n$ the following hold:

$$E(BX + b) = BE(X) + b \qquad Cov(BX + b) = BCov(X)B^T$$

Ramdom vectors (contd.)

Ramdom vectors (contd.)

Example: The d.f. f and the characteristic function ϕ_X of the multivariate normal distribution with expected value μ and covariance Σ are given as

$$f(x) = \frac{1}{\sqrt{(2\pi)^d|\Sigma|}} \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}, x \in \mathbb{R}^d$$

$$\phi_X(t) = \exp\left\{it^T\mu - \frac{1}{2}t^T\Sigma t\right\}, t \in \mathbb{R}^d,$$

where $|\Sigma| = |Det(\Sigma)|$.

Ramdom vectors (contd.)

Example: The d.f. f and the characteristic function ϕ_X of the multivariate normal distribution with expected value μ and covariance Σ are given as

$$f(x) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right\}, x \in \mathbb{R}^d$$

$$\phi_X(t) = \exp\left\{it^T\mu - \frac{1}{2}t^T \Sigma t\right\}, t \in \mathbb{R}^d,$$

where $|\Sigma| = |Det(\Sigma)|$.

Modelling the depedencies of risk factor changes (or financial data in general) in terms of the multivariate normal distribution might be inappropriate:

- risk factor changes are in general heavier tailed than normal
- the dependence between large return drops is in general stronger than the dependence between ordinary returns. This type of dependency cannot be modelled by the multivariate normal distribution.

Dependence measures

Let X_1 and X_2 be r.v. There exist several scalar measures for the dependence between X_1 und X_2.

Dependence measures

Let X_1 and X_2 be r.v. There exist several scalar measures for the dependence between X_1 und X_2.

Linear correlation

Assumption: $var(X_1), var(X_2) \in (0, \infty)$.
The linear correlation coefficient $\rho_L(X_1, X_2)$ ist given as follows

$$\rho_L(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{var(X_1) var(X_2)}}$$

Dependence measures

Let X_1 and X_2 be r.v. There exist several scalar measures for the dependence between X_1 und X_2.

Linear correlation

Assumption: $var(X_1), var(X_2) \in (0, \infty)$.
The linear correlation coefficient $\rho_L(X_1, X_2)$ ist given as follows

$$\rho_L(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{var(X_1) var(X_2)}}$$

Properties of the linear correlation coefficient:

- X_1 and X_2 are independent $\Rightarrow \rho_L(X_1, X_2) = 0$, but $\rho_L(X_1, X_2) = 0 \not\Rightarrow$ X_1 and X_2 are independent
  Example: Let $X_1 \sim N(0, 1)$ and $X_2 = X_1^2$. $\rho_L(X_1, X_2) = 0$ holds although X_1 and X_2 are dependent.

Dependence measures

Let X_1 and X_2 be r.v. There exist several scalar measures for the dependence between X_1 und X_2.

Linear correlation

Assumption: $var(X_1), var(X_2) \in (0, \infty)$.
The linear correlation coefficient $\rho_L(X_1, X_2)$ ist given as follows

$$\rho_L(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{var(X_1)var(X_2)}}$$

Properties of the linear correlation coefficient:

- X_1 and X_2 are independent $\Rightarrow \rho_L(X_1, X_2) = 0$, but $\rho_L(X_1, X_2) = 0 \not\Rightarrow$ X_1 and X_2 are independent

  Example: Let $X_1 \sim N(0, 1)$ and $X_2 = X_1^2$. $\rho_L(X_1, X_2) = 0$ holds although X_1 and X_2 are dependent.

- $|\rho_L(X_1, X_2)| = 1 \Leftrightarrow \exists \alpha, \beta \in \mathbb{R}$, $\beta \neq 0$, such that $X_2 \stackrel{d}{=} \alpha + \beta X_1$ and signum$(\beta) =$ signum$(\rho_L(X_1, X_2))$.

Properties of the linear correlation coefficient (contd.):

Properties of the linear correlation coefficient (contd.):

- The linear correlation coefficient is invariant under strict monotone increasing linear transformations. This means that for any two r.v. X_1 and X_2 and real constants $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$, $\beta_1 > 0$, $\beta_2 > 0$ the following holds:

$$\rho_L(\alpha_1 + \beta_1 X_1, \alpha_2 + \beta_2 X_2) = \rho_L(X_1, X_2).$$

Properties of the linear correlation coefficient (contd.):

- The linear correlation coefficient is invariant under strict monotone increasing linear transformations. This means that for any two r.v. X_1 and X_2 and real constants $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$, $\beta_1 > 0$, $\beta_2 > 0$ the following holds:

$$\rho_L(\alpha_1 + \beta_1 X_1, \alpha_2 + \beta_2 X_2) = \rho_L(X_1, X_2).$$

However, in general, the linear correlation coefficient is not invariant under strict monotone increasing non linear transformations.

Properties of the linear correlation coefficient (contd.):

- The linear correlation coefficient is invariant under strict monotone increasing linear transformations. This means that for any two r.v. X_1 and X_2 and real constants $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$, $\beta_1 > 0$, $\beta_2 > 0$ the following holds:

$$\rho_L(\alpha_1 + \beta_1 X_1, \alpha_2 + \beta_2 X_2) = \rho_L(X_1, X_2).$$

However, in general, the linear correlation coefficient is not invariant under strict monotone increasing non linear transformations.

Example: Let $X_1 \sim Exp(\lambda)$, $X_2 = X_1$, and T_1, T_2 be two strict monotone increasing transformations: $T_1(X_1) = X_1$ and $T_2(X_1)) = X_1^2$. Then

$$\rho_L(X_1, X_1) = 1 \text{ and } \rho_L(T_1(X_1), T_2(X_1)) = \frac{2}{\sqrt{5}}.$$

Rank correlation coefficients

Rank correlation coefficients

Let (x_1, x_2) and $(\tilde{x}_1, \tilde{x}_2)$ be two points in $\mathbb{R}^2$. They are called *concordant* iff $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$ and *discordant* iff $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$.

Rank correlation coefficients

Let (x_1, x_2) and $(\tilde{x}_1, \tilde{x}_2)$ be two points in $\mathbb{R}^2$. They are called *concordant* iff $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$ and *discordant* iff $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$.

Let $(X_1, X_2)^T$ and $(\tilde{X}_1, \tilde{X}_2)^T$ be two i.i.d. random vectors.

The Kendall's Tau ρ_τ is defined as

$$\rho_\tau(X_1, X_2) = P\left((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\right) - P\left((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0\right)$$

Rank correlation coefficients

Let (x_1, x_2) and $(\tilde{x}_1, \tilde{x}_2)$ be two points in $\mathbb{R}^2$. They are called *concordant* iff $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$ and *discordant* iff $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$.

Let $(X_1, X_2)^T$ and $(\tilde{X}_1, \tilde{X}_2)^T$ be two i.i.d. random vectors.

The Kendall's Tau ρ_τ is defined as

$$\rho_\tau(X_1, X_2) = P\Big((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\Big) - P\Big((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0\Big)$$

Let $(\hat{X}_1, \hat{X}_2)$ be a third random vector independent from (X_1, X_2) and $(\tilde{X}_1, \tilde{X}_2)$ with the same distribution as the later two vectors.

The Spearman's Rho ρ_S is defined as

$$\rho_S(X_1, X_2) = 3\Big\{P\Big((X_1 - \tilde{X}_1)(X_2 - \hat{X}_2) > 0\Big) - P\Big((X_1 - \tilde{X}_1)(X_2 - \hat{X}_2) < 0\Big)\Big\}$$

Some properties of ρ_τ und ρ_S:

Some properties of ρ_τ und ρ_S:

1. $\rho_\tau(X_1, X_2) \in [-1, 1]$ and $\rho_S(X_1, X_2) \in [-1, 1]$.

Some properties of ρ_τ und ρ_S:

1. $\rho_\tau(X_1, X_2) \in [-1, 1]$ and $\rho_S(X_1, X_2) \in [-1, 1]$.
2. if X_1 and X_2 are independent, then $\rho_\tau(X_1, X_2) = \rho_S(X_1, X_2) = 0$. In general the converse does not hold.

Some properties of ρ_τ und ρ_S:

1. $\rho_\tau(X_1, X_2) \in [-1, 1]$ and $\rho_S(X_1, X_2) \in [-1, 1]$.
2. if X_1 and X_2 are independent, then $\rho_\tau(X_1, X_2) = \rho_S(X_1, X_2) = 0$. In general the converse does not hold.
3. Let $T\colon \mathbb{R} \to \mathbb{R}$ be a strict monotone increasing function. Then the following holds

$$\rho_\tau(T(X_1), T(X_2)) = \rho_\tau(X_1, X_2)$$

$$\rho_S(T(X_1), T(X_2)) = \rho_S(X_1, X_2)$$

Proof: 1) is trivial and 2) in the case of Kendall's Tau as well.
The proof of 2) in the case of Spearman's Rho and the proof of 3) will be done in terms of copulas later.

Tail dependence

Tail dependence

Definition: Let $(X_1, X_2)^T$ be a random vector with marginal c.d.f. F_1 and F_2. *The coefficient of upper tail dependence* of $(X_1, X_2)^T$ is defined as:

$$\lambda_U(X_1, X_2) = \lim_{u \to 1^-} P(X_2 > F_2^{\leftarrow}(u) | X_1 > F_1^{\leftarrow}(u))$$

provided that this limit exists.

Tail dependence

Definition: Let $(X_1, X_2)^T$ be a random vector with marginal c.d.f. F_1 and F_2. *The coefficient of upper tail dependence* of $(X_1, X_2)^T$ is defined as:

$$\lambda_U(X_1, X_2) = \lim_{u \to 1^-} P(X_2 > F_2^{\leftarrow}(u) | X_1 > F_1^{\leftarrow}(u))$$

provided that this limit exists.

The coefficient of lower tail dependence of $(X_1, X_2)^T$ is defined as:

$$\lambda_L(X_1, X_2) = \lim_{u \to 0^+} P(X_2 \leq F_2^{\leftarrow}(u) | X_1 \leq F_1^{\leftarrow}(u))$$

provided that this limit exists.
If the limit exists and $\lambda_U > 0$ ($\lambda_L > 0$) we say that $(X_1, X_2)^T$ has an upper (lower) tail dependence.

Tail dependence

Definition: Let $(X_1, X_2)^T$ be a random vector with marginal c.d.f. F_1 and F_2. *The coefficient of upper tail dependence* of $(X_1, X_2)^T$ is defined as:

$$\lambda_U(X_1, X_2) = \lim_{u \to 1^-} P(X_2 > F_2^{\leftarrow}(u) | X_1 > F_1^{\leftarrow}(u))$$

provided that this limit exists.

The coefficient of lower tail dependence of $(X_1, X_2)^T$ is defined as:

$$\lambda_L(X_1, X_2) = \lim_{u \to 0^+} P(X_2 \leq F_2^{\leftarrow}(u) | X_1 \leq F_1^{\leftarrow}(u))$$

provided that this limit exists.
If the limit exists and $\lambda_U > 0$ ($\lambda_L > 0$) we say that $(X_1, X_2)^T$ has an upper (lower) tail dependence.

Exercise: Let $X_1 \sim Exp(\lambda)$ and $X_2 = X_1^2$. Determine $\lambda_U(X_1, X_2)$, $\lambda_L(X_1, X_2)$ and show that $(X_1, X_2)^T$ has an upper tail dependence and a lower tail dependence. Compute also the linear correlation coefficient $\rho_L(X_1, X_2)$.

Multivariate elliptical distributions

Multivariate elliptical distributions

a) The multivariate normal distribution

Definition: The random vector $(X_1, X_2, \ldots, X_d)^T$ has a *multivariate normal distribution* (or a *multivariate Gaussian distribution*) iff
$X \stackrel{d}{=} \mu + AZ$, where $Z = (Z_1, Z_2, \ldots, Z_k)^T$ is a vector of i.i.d. standard normal distributed r.v. ($Z_i \sim N(0,1)$, $\forall i = 1, 2, \ldots, k$),
$A \in \mathbb{R}^{d \times k}$ is a constant matrix and $\mu \in \mathbb{R}^d$ is a constant vector.

Multivariate elliptical distributions

a) The multivariate normal distribution

Definition: The random vector $(X_1, X_2, \ldots, X_d)^T$ has a *multivariate normal distribution* (or a *multivariate Gaussian distribution*) iff

$X \stackrel{d}{=} \mu + AZ$, where $Z = (Z_1, Z_2, \ldots, Z_k)^T$ is a vector of i.i.d. standard normal distributed r.v. ($Z_i \sim N(0,1)$, $\forall i = 1, 2, \ldots, k$), $A \in \mathbb{R}^{d\times k}$ is a constant matrix and $\mu \in \mathbb{R}^d$ is a constant vector.

For such a random vector X we have: $E(X) = \mu$, $cov(X) = \Sigma = AA^T$. Thus Σ is positive semidefinite. Notation: $X \sim N_d(\mu, \Sigma)$.

Multivariate elliptical distributions

a) The multivariate normal distribution

Definition: The random vector $(X_1, X_2, \ldots, X_d)^T$ has a *multivariate normal distribution* (or a *multivariate Gaussian distribution*) iff
$X \stackrel{d}{=} \mu + AZ$, where $Z = (Z_1, Z_2, \ldots, Z_k)^T$ is a vector of i.i.d. standard normal distributed r.v. ($Z_i \sim N(0,1)$, $\forall i = 1, 2, \ldots, k$),
$A \in \mathbb{R}^{d\times k}$ is a constant matrix and $\mu \in \mathbb{R}^d$ is a constant vector.

For such a random vector X we have: $E(X) = \mu$, $cov(X) = \Sigma = AA^T$.
Thus Σ is positive semidefinite. Notation: $X \sim N_d(\mu, \Sigma)$.

Theorem: (Equivalent characterisations of the multivariate normal distribution)

1. $X \sim N_d(\mu, \Sigma)$ for some vector $\mu \in \mathbb{R}^d$ and some positive semidefinite matrix $\Sigma \in \mathbb{R}^{d\times d}$, iff $\forall a \in \mathbb{R}^d$, $a = (a_1, a_2, \ldots, a_d)^T$, the random variable $a^T X$ is normally distributed.

Equivalent characterisations of the multivariate normal distribution

2. A random vector $X \in \mathbb{R}^d$ is multivariate normally distributed iff its characteristic function $\phi_X(t)$ is given as

$$\phi_X(t) = E(\exp\{it^T X\}) = \exp\{it^T \mu - \frac{1}{2} t^T \Sigma t\}$$

for some vector $\mu \in \mathbb{R}^d$ and some positive semidefinite matrix $\Sigma \in \mathbb{R}^{d \times d}$.

3. A random vector $X \in \mathbb{R}^d$ with $E(X) = \mu$ and $cov(X) = \Sigma$, $|\Sigma| > 0$, is multivariate normally distributed, i.e. $X \sim N_d(\mu, \Sigma)$, iff its density function $f_X(x)$ is given as follows

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left\{ -\frac{(x-\mu)^T \Sigma^{-1} (x-\mu)}{2} \right\}.$$

Proof: (see eg. Gut 1995)

Properties of the multivariate normal distribution

Properties of the multivariate normal distribution

Theorem:

Let $X \sim N_d(\mu, \Sigma)$. The following hold:

- Linear combinations:
  Let $B \in \mathbb{R}^{k \times d}$ and $b \in \mathbb{R}^k$. Then $BX + b \in N_k(B\mu + b, B\Sigma B^T)$.

Properties of the multivariate normal distribution

Theorem:

Let $X \sim N_d(\mu, \Sigma)$. The following hold:

- Linear combinations:
  Let $B \in \mathbb{R}^{k \times d}$ and $b \in \mathbb{R}^k$. Then $BX + b \in N_k(B\mu + b, B\Sigma B^T)$.

- Marginal distributions:
  Let $X^T = \left({X^{(1)}}^T, {X^{(2)}}^T \right)$ with ${X^{(1)}}^T = (X_1, X_2, \ldots, X_k)^T$ and ${X^{(2)}}^T = (X_{k+1}, X_{k+2}, \ldots, X_d)^T$. Analogously let

$$\mu^T = \left({\mu^{(1)}}^T, {\mu^{(2)}}^T \right) \text{ and } \Sigma = \begin{pmatrix} \Sigma^{(1,1)} & \Sigma^{(1,2)} \\ \Sigma^{(2,1)} & \Sigma^{(2,2)} \end{pmatrix}.$$

  Then $X^{(1)} \sim N_k\left(\mu^{(1)}, \Sigma^{(1,1)}\right)$ and $X^{(2)} \sim N_{d-k}\left(\mu^{(2)}, \Sigma^{(2,2)}\right)$.

Properties of the multivariate normal distribution (contd.)

Properties of the multivariate normal distribution (contd.)

- Conditional distributions:

  Let Σ be nonsingular. The conditioned random vector $X^{(2)} \Big| X^{(1)} = x^{(1)}$ is multivariate normally distributed with

$$X^{(2)}|X^{(1)} = x^{(1)} \sim N_{d-k}\left(\mu^{(2,1)}, \Sigma^{(22,1)}\right) \text{ where}$$

$$\mu^{(2,1)} = \mu^{(2)} + \Sigma^{(2,1)}\left(\Sigma^{(1,1)}\right)^{-1}\left(x^{(1)} - \mu^{(1)}\right) \text{ and}$$

$$\Sigma^{(22,1)} = \Sigma^{(2,2)} - \Sigma^{(2,1)}\left(\Sigma^{(1,1)}\right)^{-1}\Sigma^{(1,2)}.$$

Properties of the multivariate normal distribution (contd.)

Properties of the multivariate normal distribution (contd.)

- Quadratic forms:

  Is Σ is nonsingular, then $D^2 = (X - \mu)^T \Sigma^{-1} (X - \mu) \sim \chi^2_d$. The r.v. D is called *Mahalanobis distance*.

Properties of the multivariate normal distribution (contd.)

- Quadratic forms:

  Is Σ is nonsingular, then $D^2 = (X - \mu)^T \Sigma^{-1} (X - \mu) \sim \chi^2_d$. The r.v. D is called *Mahalanobis distance*.

- Convolutions:

  Let $X \sim N_d(\mu, \Sigma)$ and $Y \sim N_d(\tilde{\mu}, \tilde{\Sigma})$ be two independent random vectors. Then $X + Y \sim N_d(\mu + \tilde{\mu}, \Sigma + \tilde{\Sigma})$.

Normal mixture

Normal mixture

Definition: A random vector $X \in \mathbb{R}^d$ is said to have a multivariate normal variance mixture distribution if $X \stackrel{d}{=} \mu + WAZ$ where $Z \sim N_k(0, I)$, $W \geq 0$ is a r.v. independent from Z, $\mu \in \mathbb{R}^d$ is a constant vector, $A \in \mathbb{R}^{d \times k}$ is a constant matrix, and I is the unit matrix.

Normal mixture

Definition: A random vector $X \in \mathbb{R}^d$ is said to have a multivariate normal variance mixture distribution if $X \stackrel{d}{=} \mu + WAZ$ where $Z \sim N_k(0, I)$, $W \geq 0$ is a r.v. independent from Z, $\mu \in \mathbb{R}^d$ is a constant vector, $A \in \mathbb{R}^{d \times k}$ is a constant matrix, and I is the unit matrix.

By conditioning on $W = w$ we get $X \sim N_d(\mu, w^2\Sigma)$, with $\Sigma = AA^T$.

Normal mixture

Definition: A random vector $X \in \mathbb{R}^d$ is said to have a multivariate normal variance mixture distribution if $X \stackrel{d}{=} \mu + WAZ$ where $Z \sim N_k(0, I)$, $W \geq 0$ is a r.v. independent from Z, $\mu \in \mathbb{R}^d$ is a constant vector, $A \in \mathbb{R}^{d \times k}$ is a constant matrix, and I is the unit matrix.

By conditioning on $W = w$ we get $X \sim N_d(\mu, w^2\Sigma)$, with $\Sigma = AA^T$.

Moreover $E(X) = \mu$ and $cov(X) = E(W^2 AZZ^T A^T) = E(W^2)\Sigma$, if $E(W^2) < \infty$

Normal mixture

Definition: A random vector $X \in \mathbb{R}^d$ is said to have a multivariate normal variance mixture distribution if $X \stackrel{d}{=} \mu + WAZ$ where $Z \sim N_k(0, I)$, $W \geq 0$ is a r.v. independent from Z, $\mu \in \mathbb{R}^d$ is a constant vector, $A \in \mathbb{R}^{d \times k}$ is a constant matrix, and I is the unit matrix.

By conditioning on $W = w$ we get $X \sim N_d(\mu, w^2\Sigma)$, with $\Sigma = AA^T$.

Moreover $E(X) = \mu$ and $cov(X) = E(W^2 AZZ^T A^T) = E(W^2)\Sigma$, if $E(W^2) < \infty$

Example: the multivariate t_α distribution
Let $Y \sim IG(\alpha, \beta)$ (inverse-gamma distribution) with density function given as $f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} \exp(-\beta/x)$ for $x > 0$, $\alpha > 0$, $\beta > 0$.

Then $E(Y) = \frac{\beta}{\alpha-1}$ for $\alpha > 1$, $var(Y) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$.

Normal mixture

Definition: A random vector $X \in \mathbb{R}^d$ is said to have a multivariate normal variance mixture distribution if $X \stackrel{d}{=} \mu + WAZ$ where $Z \sim N_k(0, I)$, $W \geq 0$ is a r.v. independent from Z, $\mu \in \mathbb{R}^d$ is a constant vector, $A \in \mathbb{R}^{d\times k}$ is a constant matrix, and I is the unit matrix.

By conditioning on $W = w$ we get $X \sim N_d(\mu, w^2\Sigma)$, with $\Sigma = AA^T$.

Moreover $E(X) = \mu$ and $cov(X) = E(W^2 AZZ^T A^T) = E(W^2)\Sigma$, if $E(W^2) < \infty$

Example: the multivariate t_α distribution

Let $Y \sim IG(\alpha, \beta)$ (inverse-gamma distribution) with density function given as $f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} \exp(-\beta/x)$ for $x > 0$, $\alpha > 0$, $\beta > 0$.

Then $E(Y) = \frac{\beta}{\alpha-1}$ for $\alpha > 1$, $var(Y) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$.

Let $W^2 \sim IG(\alpha/2, \alpha/2)$. Then $X = \mu + WAZ$ has *the multivariate t_α-distribution with α degrees of freedom*. Notation: $X \sim t_d(\alpha, \mu, \Sigma)$.

Normal mixture

Definition: A random vector $X \in \mathbb{R}^d$ is said to have a multivariate normal variance mixture distribution if $X \stackrel{d}{=} \mu + WAZ$ where $Z \sim N_k(0, I)$, $W \geq 0$ is a r.v. independent from Z, $\mu \in \mathbb{R}^d$ is a constant vector, $A \in \mathbb{R}^{d \times k}$ is a constant matrix, and I is the unit matrix.

By conditioning on $W = w$ we get $X \sim N_d(\mu, w^2\Sigma)$, with $\Sigma = AA^T$.

Moreover $E(X) = \mu$ and $cov(X) = E(W^2 AZZ^T A^T) = E(W^2)\Sigma$, if $E(W^2) < \infty$

Example: the multivariate t_α distribution

Let $Y \sim IG(\alpha, \beta)$ (inverse-gamma distribution) with density function given as $f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} \exp(-\beta/x)$ for $x > 0$, $\alpha > 0$, $\beta > 0$.

Then $E(Y) = \frac{\beta}{\alpha - 1}$ for $\alpha > 1$, $var(Y) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ for $\alpha > 2$.

Let $W^2 \sim IG(\alpha/2, \alpha/2)$. Then $X = \mu + WAZ$ has *the multivariate t_α-distribution with α degrees of freedom*. Notation: $X \sim t_d(\alpha, \mu, \Sigma)$.

Since $E(W^2) = \alpha/(\alpha - 2)$, for $\alpha > 2$, we get $cov(X) = E(W^2)\Sigma = \frac{\alpha}{\alpha-2}\Sigma$.

Spherical distributions

Spherical distributions

Definition: A random vector $X = (X_1, X_2, \ldots, X_d)^T$ has a spherical distribution if for every orthogonal matrix $U \in \mathbb{R}^{d\times d}$ we have $UX \stackrel{d}{=} X$.

Spherical distributions

Definition: A random vector $X = (X_1, X_2, \ldots, X_d)^T$ has a spherical distribution if for every orthogonal matrix $U \in \mathbb{R}^{d \times d}$ we have $UX \stackrel{d}{=} X$.

Theorem: The following statements are equivalent:

1. $X \in \mathbb{R}^d$ has a spherical distribution.

Spherical distributions

Definition: A random vector $X = (X_1, X_2, \ldots, X_d)^T$ has a spherical distribution if for every orthogonal matrix $U \in \mathbb{R}^{d \times d}$ we have $UX \stackrel{d}{=} X$.

Theorem: The following statements are equivalent:

1. $X \in \mathbb{R}^d$ has a spherical distribution.
2. There exists a function $\psi \colon \mathbb{R} \to \mathbb{R}$ of a scalar variable, such that the characteristic function of X satisfies

$$\phi_X(t) = \psi(t^T t) = \psi(t_1^2 + t_2^2 + \ldots + t_d^2)$$

Spherical distributions

Definition: A random vector $X = (X_1, X_2, \ldots, X_d)^T$ has a spherical distribution if for every orthogonal matrix $U \in \mathbb{R}^{d \times d}$ we have $UX \stackrel{d}{=} X$.

Theorem: The following statements are equivalent:

1. $X \in \mathbb{R}^d$ has a spherical distribution.
2. There exists a function $\psi \colon \mathbb{R} \to \mathbb{R}$ of a scalar variable, such that the characteristic function of X satisfies

$$\phi_X(t) = \psi(t^T t) = \psi(t_1^2 + t_2^2 + \ldots + t_d^2)$$

3. For every vector $a \in \mathbb{R}^d$, $a^t X \stackrel{d}{=} ||a|| X_1$ holds, where $||a||^2 = a_1^2 + a_2^2 + \ldots + a_d^2$.

Spherical distributions

Definition: A random vector $X = (X_1, X_2, \ldots, X_d)^T$ has a spherical distribution if for every orthogonal matrix $U \in \mathbb{R}^{d \times d}$ we have $UX \stackrel{d}{=} X$.

Theorem: The following statements are equivalent:

1. $X \in \mathbb{R}^d$ has a spherical distribution.
2. There exists a function $\psi \colon \mathbb{R} \to \mathbb{R}$ of a scalar variable, such that the characteristic function of X satisfies
$$\phi_X(t) = \psi(t^T t) = \psi(t_1^2 + t_2^2 + \ldots + t_d^2)$$
3. For every vector $a \in \mathbb{R}^d$, $a^t X \stackrel{d}{=} ||a|| X_1$ holds, where $||a||^2 = a_1^2 + a_2^2 + \ldots + a_d^2$.
4. X has the stochastic representation $X \stackrel{d}{=} RS$, where $S \in \mathbb{R}^d$ is a random vector uniformly distributed on the unit sphere S^{d-1}, $\mathcal{S}^{d-1} := \{x \in \mathbb{R}^d \colon ||x|| = 1\}$, and $R \geq 0$ is a r.v. independent of S.

Notation: $X \sim S_d(\psi)$, cf. 2.