IEOR E4602: Quantitative Risk Management

Multivariate Distributions

Martin Haugh

Department of Industrial Engineering and Operations Research
Columbia University
Email: martin.b.haugh@gmail.com

Joint and Marginal CDFs

Let $\mathbf{X} = (X_1, \ldots X_n)$ is an n-dimensional vector of random variables.

Definition (Joint CDF): For all $\mathbf{x} = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$, the joint cumulative distribution function (CDF) of $\mathbf{X}$ satisfies

$$F_{\mathbf{X}}(\mathbf{x}) \;=\; F_{\mathbf{X}}(x_1, \ldots, x_n) \;=\; P(X_1 \leq x_1, \ldots, X_n \leq x_n).$$

Definition (Marginal CDF): For a fixed i, the marginal CDF of X_i satisfies

$$F_{X_i}(x_i) \;=\; F_{\mathbf{X}}(\infty, \ldots, \infty, x_i, \infty, \ldots \infty).$$

Straightforward to generalize to joint marginal distributions. e.g

$$F_{ij}(x_i, x_j) \;=\; F_{\mathbf{X}}(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty, x_j, \infty, \ldots \infty).$$

Conditional CDFs

If $\mathbf{X}$ has a probability density function (PDF) then

$$F_{\mathbf{X}}(x_1, \ldots, x_n) \;=\; \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(u_1, \ldots, u_n)\; du_1 \ldots du_n.$$

A collection of random variables is independent if the joint CDF (or PDF if it exists) can be factored into the product of the marginal CDFs (or PDFs).

If $\mathbf{X_1} = (X_1, \ldots, X_k)^\top$ and $\mathbf{X_2} = (X_{k+1}, \ldots, X_n)^\top$ is a *partition* of $\mathbf{X}$ then the conditional CDF satisfies

$$F_{\mathbf{X_2}|\mathbf{X_1}}(\mathbf{x_2}|\mathbf{x_1}) \;=\; P(\mathbf{X_2} \leq \mathbf{x_2}|\mathbf{X_1} = \mathbf{x_1}).$$

If $\mathbf{X}$ has a PDF, $f(\cdot)$, then it satisfies

$$F_{\mathbf{X_2}|\mathbf{X_1}}(\mathbf{x_2}|\mathbf{x_1}) \;=\; \int_{-\infty}^{x_{k+1}} \cdots \int_{-\infty}^{x_n} \frac{f(x_1, \ldots, x_k, u_{k+1}, \ldots, u_n)}{f_{\mathbf{X_1}}(\mathbf{x_1})}\; du_{k+1} \ldots du_n$$

where $f_{\mathbf{X_1}}(\cdot)$ is the joint marginal PDF of $\mathbf{X_1}$.

Mean Vector and Covariance Matrix

Assuming it exists, mean vector of $\mathbf{X}$ given by

$$\mathsf{E}[\mathbf{X}] := (\mathsf{E}[X_1] \; \ldots \; \mathsf{E}[X_n])^\top .$$

Again assuming it exists, the covariance matrix of $\mathbf{X}$ satisfies

$$\mathsf{Cov}(\mathbf{X}) := \mathbf{\Sigma} := \mathsf{E}\left[(\mathbf{X} - \mathsf{E}[\mathbf{X}])\,(\mathbf{X} - \mathsf{E}[\mathbf{X}])^\top\right]$$

so that the $(i,j)^{th}$ element of Σ is simply the covariance of X_i and X_j.

Important properties of $\mathbf{\Sigma}$:

1. It is symmetric so that $\mathbf{\Sigma}^\top = \mathbf{\Sigma}$
2. Diagonal elements satisfy $\mathbf{\Sigma}_{i,i} \geq 0$
3. It is positive semi-definite so that $x^\top \mathbf{\Sigma} x \geq 0$ for all $x \in \mathbb{R}^n$.

The correlation matrix, $\rho(\mathbf{X})$, has $(i,j)^{th}$ element $\rho_{ij} := \mathsf{Corr}(X_i, X_j)$

- also symmetric, positive semi-definite
- has 1's along the diagonal.

Linear Combinations and Characteristic Functions

For any matrix $\mathbf{A} \in \mathbb{R}^{k \times n}$ and vector $\mathbf{a} \in \mathbb{R}^k$ have

$$\mathsf{E}[\mathbf{AX} + \mathbf{a}] = \mathbf{A}\mathsf{E}[\mathbf{X}] + \mathbf{a} \tag{1}$$

$$\mathsf{Cov}(\mathbf{AX} + \mathbf{a}) = \mathbf{A} \, \mathsf{Cov}(\mathbf{X}) \, \mathbf{A}^\top. \tag{2}$$

The characteristic function of $\mathbf{X}$ given by

$$\phi_{\mathbf{X}}(s) := \mathsf{E}\left[e^{is^\top \mathbf{X}}\right] \quad \text{for} \quad s \in \mathbb{R}^n \tag{3}$$

If it exists, the moment-generating function (MGF) is given by (3) with s replaced by $-is$.

The Multivariate Normal Distribution

If $\mathbf{X}$ multivariate normal with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ then write

$$\mathbf{X} \sim \mathsf{MN}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}).$$

Standard multivariate normal: $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{I_n}$, the $n \times n$ identity matrix.

PDF of $\mathbf{X}$ given by

$$f(\mathbf{x}) \;=\; \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}}\; e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\top\, \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} \tag{4}$$

where $|\cdot|$ denotes the determinant.

Characteristic function satisfies

$$\phi_{\mathbf{X}}(s) \;=\; \mathsf{E}\left[e^{is^\top \mathbf{X}}\right] \;=\; e^{is^\top \boldsymbol{\mu} \,-\, \frac{1}{2}s^\top \boldsymbol{\Sigma} s}.$$

The Multivariate Normal Distribution

Let $\mathbf{X_1} = (X_1, \ldots, X_k)^\top$ and $\mathbf{X_2} = (X_{k+1}, \ldots, X_n)^\top$ be a partition of $\mathbf{X}$ with

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix} \quad \text{and} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

Then marginal distribution of a multivariate normal random vector is itself (multivariate) normal. In particular, $\mathbf{X_i} \sim \mathsf{MN}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii})$, for $i = 1, 2$.

Assuming $\boldsymbol{\Sigma}$ is positive definite, the conditional distribution of a multivariate normal distribution is also a (multivariate) normal distribution. In particular,

$$\mathbf{X_2} \mid \mathbf{X_1} = \mathbf{x_1} \ \sim \ \mathsf{MN}(\boldsymbol{\mu}_{\mathbf{2.1}}, \boldsymbol{\Sigma}_{2.1})$$

where

$$\boldsymbol{\mu}_{2.1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\, \boldsymbol{\Sigma}_{11}^{-1}\, (\mathbf{x_1} - \boldsymbol{\mu}_1)$$

$$\boldsymbol{\Sigma}_{2.1} = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}.$$

Generating MN Distributed Random Vectors

Suppose we wish to generate $\mathbf{X} = (X_1, \ldots, X_n)$ where $\mathbf{X} \sim \mathsf{MN}_n(\mathbf{0}, \mathbf{\Sigma})$

- it is then easy to handle the case where $\mathsf{E}[\mathbf{X}] \neq \mathbf{0}$.

Let $\mathbf{Z} = (Z_1, \ldots, Z_n)^\top$ where $Z_i \sim \mathsf{N}(0,1)$ and IID for $i = 1, \ldots, n$.

If $\mathbf{C}$ an $(n \times m)$ matrix then

$$\mathbf{C}^\top \mathbf{Z} \sim \mathsf{MN}(0, \mathbf{C}^\top \mathbf{C}).$$

Problem therefore reduces to finding $\mathbf{C}$ such that $\mathbf{C}^\top \mathbf{C} = \mathbf{\Sigma}$.

Usually find such a matrix, $\mathbf{C}$, via the Cholesky decomposition of $\mathbf{\Sigma}$.

The Cholesky Decomposition of a Symmetric PD Matrix

Any symmetric positive-definite matrix, $\mathbf{M}$, may be written as

$$\mathbf{M} = \mathbf{U}^\top \mathbf{D} \mathbf{U}$$

where:

- $\mathbf{U}$ is an upper triangular matrix
- $\mathbf{D}$ is a diagonal matrix with positive diagonal elements.

Since $\mathbf{\Sigma}$ is symmetric positive-definite, can therefore write

$$\begin{aligned} \mathbf{\Sigma} &= \mathbf{U}^\top \mathbf{D} \mathbf{U} \\ &= (\mathbf{U}^\top \sqrt{\mathbf{D}})(\sqrt{\mathbf{D}}\mathbf{U}) \\ &= (\sqrt{\mathbf{D}}\mathbf{U})^\top (\sqrt{\mathbf{D}}\mathbf{U}). \end{aligned}$$

$\mathbf{C} = \sqrt{\mathbf{D}}\mathbf{U}$ therefore satisfies $\mathbf{C}^\top \mathbf{C} = \mathbf{\Sigma}$

- **C** is called the Cholesky Decomposition of $\mathbf{\Sigma}$.

The Cholesky Decomposition in Matlab

Easy to compute the Cholesky decomposition of a symmetric positive-definite matrix in Matlab using the chol command

- so also easy to simulate multivariate normal random vectors in Matlab.

Sample Matlab Code

```
>> Sigma = [1.0 0.5 0.5;
            0.5 2.0 0.3;
            0.5 0.3 1.5];

>> C = chol(Sigma);
>> Z = randn(3,1000000);
>> X = C'*Z;
>> cov(X')

ans =
    0.9972    0.4969    0.4988
    0.4969    1.9999    0.2998
    0.4988    0.2998    1.4971
```

The Cholesky Decomposition in `Matlab` and `R`

Must be very careful in `Matlab` and `R` to pre-multiply $\mathbf{Z}$ by $\mathbf{C}^\top$ and not $\mathbf{C}$.

Some languages take $\mathbf{C}^\top$ to be the Cholesky Decomposition rather $\mathbf{C}$

- must therefore always know what convention your programming language / package is using.

Must also be careful that $\mathbf{\Sigma}$ is indeed a genuine variance-covariance matrix.

Normal-Mixture Models

Normal-mixture models are a class of models generated by introducing randomness into the covariance matrix and / or the mean vector:

Definition: The random vector $\mathbf{X}$ has a normal variance mixture if

$$\mathbf{X} \sim \boldsymbol{\mu} \; + \; \sqrt{W}\, \mathbf{A}\mathbf{Z}$$

where

(i) $\mathbf{Z} \sim \mathsf{MN}_k(\mathbf{0}, \mathbf{I_k})$

(ii) $W \geq 0$ is a scalar random variable independent of $\mathbf{Z}$ and

(iii) $\mathbf{A} \in \mathbb{R}^{n \times k}$ and $\boldsymbol{\mu} \in \mathbb{R}^n$ are a matrix and vector of constants, respectively.

Normal-Mixture Models

If we condition on W, then $\mathbf{X}$ is multivariate normally distributed

- this observation also leads to an obvious simulation algorithm for generating samples of $\mathbf{X}$.

Typically interested in case when $\text{rank}(A) = n \leq k$ and $\mathbf{\Sigma}$ is a full-rank positive definite matrix

- then obtain a non-singular normal variance mixture.

Assuming W is integrable, immediately see that

$$\text{E}[\mathbf{X}] \;=\; \boldsymbol{\mu} \quad \text{and} \quad \text{Cov}(\mathbf{X}) \;=\; \text{E}[W]\,\mathbf{\Sigma}$$

where $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^\top$.

We call $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$ the location vector and dispersion matrix of the distribution.

Also clear that correlation matrices of $\mathbf{X}$ and $\mathbf{AZ}$ coincide

- implies that if $\mathbf{A} = \mathbf{I}_n$ then components of $\mathbf{X}$ are uncorrelated though they are **not** in general independent.

Normal-Mixture Models

Lemma: Let $\mathbf{X} = (X_1, X_2)$ have a normal mixture distribution with $\mathbf{A} = \mathbf{I_2}$, $\boldsymbol{\mu} = \mathbf{0}$ and $\mathsf{E}[W] < \infty$ so that $\mathrm{Cov}(X_1, X_2) = 0$.

Then X_1 and X_2 are independent if and only if W is constant with probability 1.

Proof: (i) If W constant then immediately follows from independence of Z_1 and Z_2 that X_1 and X_2 are also independent.

(ii) Suppose now X_1 and X_2 are independent. Note that

$$
\begin{aligned}
\mathsf{E}[|X_1|\,|X_2|] = \mathsf{E}[W\,|Z_1|\,|Z_2|] &= \mathsf{E}[W]\,\mathsf{E}[|Z_1|\,|Z_2|] \\
&\geq \left(\mathsf{E}[\sqrt{W}]\right)^2\,\mathsf{E}[|Z_1|\,|Z_2|] \\
&= \mathsf{E}[|X_1|]\,\mathsf{E}[|X_2|]
\end{aligned}
$$

with equality only if W is a constant.

But independence of X_1 and X_2 implies we must have equality and so W is indeed constant almost surely. $\square$

E.G. The Multivariate Two-Point Normal Mixture Model

Perhaps the simplest example of normal-variance mixture is obtained when W is a discrete random variable.

If W is binary and takes on two values, w_1 and w_2 with probabilities p and $1-p$, respectively, then obtain the two-point normal mixture model.

Can create a two regime model by setting w_2 large relative to w_1 and choosing p large

- then $W = w_1$ can correspond to an ordinary regime
- and $W = w_2$ corresponds to a stress regime.

E.G. The Multivariate t Distribution

The multivariate t distribution with ν degrees-of-freedom (dof) is obtained when we take W to have an inverse gamma distribution.

Equivalently, the multivariate t distribution with ν dof is obtained if $\nu/W \sim \chi^2_\nu$

- the more familiar description of the t distribution.

We write $X \sim t_n(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Note that $\text{Cov}(\mathbf{X}) \;=\; \nu/(\nu-2)\boldsymbol{\Sigma}$

- only defined when $\nu > 2$.

Can easily simulate chi-squared random variables so easy to simulate multivariate t random vectors.

The multivariate t distribution plays an important role in risk management as it often provides a very good fit to asset return distributions.

Characteristic Function of a Normal Variance Mixture

We have

$$\begin{aligned}\phi_{\mathbf{X}}(s) \;=\; \mathsf{E}\left[e^{is^\top \mathbf{X}}\right] &= \mathsf{E}\left[\mathsf{E}\left[e^{is^\top \mathbf{X}} \,|\, W\right]\right] \\ &= \mathsf{E}\left[e^{is^\top \boldsymbol{\mu} \,-\, \frac{1}{2} W s^\top \boldsymbol{\Sigma} s}\right] \\ &= e^{is^\top \boldsymbol{\mu}} \;\hat{W}\left(\frac{1}{2} s^\top \boldsymbol{\Sigma} s\right)\end{aligned}$$

where $\hat{W}(\cdot)$ is the Laplace transform of W.

Sometimes use the notation $\mathbf{X} \sim \mathsf{M}_n\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \hat{W}\right)$.

Affine Transformations of Normal Variance Mixtures

Proposition: If $\mathbf{X} \sim \mathsf{M}_n\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \hat{W}\right)$ and $\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{b}$ for $\mathbf{B} \in \mathbb{R}^{k \times n}$ and $\mathbf{b} \in \mathbb{R}^k$ then $\mathbf{Y} \sim \mathsf{M}_k\left(\mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^\top, \hat{W}\right)$.

So affine transformations of normal variance mixtures remain normal variance mixtures

- useful when loss function is approximated with linear function of risk factors.

Proof is straightforward using characteristic function argument.

Normal Mean-Variance Mixtures

Could also define normal mixture distributions where $\boldsymbol{\mu} = m(W)$.

Would still obtain that $\mathbf{X}$ is multivariate normal conditional on W.

Important class of normal mean-variance mixtures are the generalized hyperbolic distributions. They:

- are closed under addition
- are easy to simulate
- can be fitted using standard statistical techniques.

We will not study normal mean-variance mixtures in this course.

Spherical Distributions

Recall that a linear transformation $\mathbf{U} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}$ is orthogonal if $\mathbf{U}\mathbf{U}^\top = \mathbf{U}^\top\mathbf{U} = \mathbf{I}_n$.

Definition: A random vector $\mathbf{X} = (X_1, \ldots, X_n)$ has a spherical distribution if

$$\mathbf{U}\mathbf{X} \;\sim\; \mathbf{X} \tag{5}$$

for every orthogonal linear transformation, $\mathbf{U} \in \mathbb{R}^{\mathbf{n}\times\mathbf{n}}$.

Note that (5) implies the distribution of $\mathbf{X}$ is **invariant** under rotations.

A better understanding of spherical distributions may be obtained from the following theorem ....

Spherical Distributions

Theorem: The following are equivalent:

1. $\mathbf{X}$ is spherical.
2. There exists a function $\psi(\cdot)$ such that for all $\mathbf{s} \in \mathbb{R}^n$,

$$\phi_{\mathbf{X}}(\mathbf{s}) = \psi(\mathbf{s}^\top \mathbf{s}) = \psi(s_1^2 + \cdots + s_n^2). \tag{6}$$

3. For all $\mathbf{a} \in \mathbb{R}^n$

$$\mathbf{a}^\top X \sim ||\mathbf{a}|| \, X_1$$

   where $||\mathbf{a}||^2 = \mathbf{a}^\top \mathbf{a} = a_1^2 + \cdots + a_n^2$.

(6) shows that characteristic function of a spherical distribution is completely determined by a function, $\psi(\cdot)$, of a scalar variable.

$\psi(\cdot)$ is known as the generator of the distribution

- common to write $\mathbf{X} \sim \mathsf{S}_n(\psi)$.

Example: Multivariate Normal

Let $\mathbf{X} \sim \mathsf{MN}_n(\mathbf{0}, \mathbf{I}_n)$. Then

$$\phi_{\mathbf{X}}(\mathbf{s}) \;=\; e^{-\frac{1}{2}\mathbf{s}^\top \mathbf{s}}.$$

So $\mathbf{X}$ is spherical with generator $\psi(s) = \exp(-\frac{1}{2}\,s)$.

Example: Normal Variance Mixtures

Suppose $\mathbf{X} \sim \mathsf{M}_n\left(\mathbf{0}, \mathbf{I}_n, \hat{W}\right)$

- so $\mathbf{X}$ has a standardized, uncorrelated normal variance mixture.

Then part (2) of previous theorem implies that $\mathbf{X}$ is spherical with $\psi(s) = \hat{W}(s/2)$.

Note there are spherical distributions that are **not** normal variance mixture distributions.

Now for another important and insightful result . . .

Spherical Distributions

Theorem: The random vector $\mathbf{X} = (X_1, \ldots, X_n)$ has a spherical distribution if and only if it has the representation

$$\mathbf{X} \sim R\,\mathbf{S}$$

where:

1. $\mathbf{S}$ is uniformly distributed on the unit sphere: $\mathcal{S}^{n-1} := \{\mathbf{s} \in \mathbb{R}^n \ : \ \mathbf{s}^\top \mathbf{s} = 1\}$ and
2. $R \geq 0$ is a random variable independent of $\mathbf{S}$.

Elliptical Distributions

Definition: The random vector $\mathbf{X} = (X_1, \ldots X_n)$ has an elliptical distribution if

$$\mathbf{X} \sim \boldsymbol{\mu} + \mathbf{A}\,\mathbf{Y}$$

where $\mathbf{Y} \sim \mathsf{S}_k(\psi)$ and $\mathbf{A} \in \mathbb{R}^{n \times k}$ and $\boldsymbol{\mu} \in \mathbb{R}^n$ are a matrix and vector of constants, respectively.

Elliptical distributions therefore obtained via multivariate affine transformations of spherical distributions.

Characteristic Function of Elliptical Distributions

Easy to calculate characteristic function of an elliptical distribution:

$$\begin{aligned}\phi_{\mathbf{X}}(\mathbf{s}) &= \mathsf{E}\left[e^{i\mathbf{s}^\top(\boldsymbol{\mu}+\mathbf{A}\ \mathbf{Y})}\right] \\ &= e^{i\mathbf{s}^\top\boldsymbol{\mu}}\ \mathsf{E}\left[e^{i(\mathbf{A}^\top\mathbf{s})^\top\mathbf{Y}}\right] \\ &= e^{i\mathbf{s}^\top\boldsymbol{\mu}}\ \psi\left(\mathbf{s}^\top\boldsymbol{\Sigma}\mathbf{s}\right)\end{aligned}$$

where as before $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^\top$.

Common to write $\mathbf{X} \sim \mathsf{E}_n\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi\right)$

- $\boldsymbol{\mu}$ known as the location vector
- $\boldsymbol{\Sigma}$ known as the dispersion matrix.

But $\boldsymbol{\Sigma}$ and ψ only uniquely determined up to a positive constant.