Multivariate Statistics (Week 1): Introduction and random vectors

Erwan Koch

EPFL (Institute of Mathematics)

Technicalities

- Lecture: Thursday 13:15–15:00.
- Exercices: Thursday 15:15–17:00.
- Teaching assistants: Sonia Alouini (sonia.alouini@epfl.ch), Timmy Tse (timmy.tse@epfl.ch)
- All the material (course description, slides, exercises, solutions and data) will be on Moodle.
- Evaluation: written exam (3 hours).
- Go to exercise sessions, it will help you a lot!
- Work in groups.
- Everyone in the class should ask at least two questions during each lecture.

Outline

Outline

Aim of multivariate analysis

Aim of multivariate statistics

Study of data containing observations on two or more “variables” measured on a set of “objects”.

Examples of such data

- Data: Grades obtained at different exams by several students, flower features for different species of iris, rainfall amount at different stations at different time points, price of different financial assets at different time points.
- Variables: grade at each exam (e.g., statistics), each flower feature, rainfall amount at each station, price of each financial asset.
- Objects: students, flowers, time points, time points.

Terminology

- Synonyms of “variable": attribute, characteristic, description, response
- Synonyms of “object”: individual, observation.

Depends on the context.

Data matrix

- We denote by n the number of objects and p the number of variables. The np pieces of information generally arranged in a $(n \times p)$ data matrix.
- The general $(n \times p)$ data matrix with n objects and p variables is

$$X = \begin{pmatrix} X_{11} & \cdots & X_{1p} \\ X_{21} & \ddots & X_{2p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{np} \end{pmatrix}$$

- Each row corresponds to an object and each column to a variable.
- We have n observations of a generic random vector (rv) $\mathbf{X} = (X_1, \ldots, X_p)'$ containing the variables $X_1, \ldots, X_p$.
  $\Longrightarrow$ Multivariate statistics deal with observations of (possibly highly-dimensional) rvs.

Types of data

- Quantitative variables:
  - Continuous variables.
  - Discrete variables.
- Qualitative (categorical) variables:
  - Ordinal variables: can be ordered in a natural way (e.g., your interest for the course).
  - Nominal variables: cannot be ordered in a meaningful way (e.g., gender, blood group, tree type).
- Quantitative and qualitative variables are sometimes said to be metric and non-metric, respectively.
- The variables studied need not all be of the same type.

Data set example

Mechanics (C)	Physics (C)	Algebra (O)	Analysis (O)	Statistics (O)	Age	Gender
77	82	67	67	81	22	F
63	78	80	70	81	23	M
75	73	71	66	81	21	F
55	72	63	70	68	22	M
63	63	65	70	63	23	M
53	61	72	64	73	22	F
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Grades (out of 100) of several students at different exams, age and gender. (C) indicates closed-book and (O) indicates open-book.

Multivariate problems and methods

- Necessary to have at our disposal multivariate statistical models (e.g., multivariate distributions). An important aspect: analysis of the (extremal) dependence between different variables (e.g., rainfall amount at several sites, returns of different financial assets) using multivariate models. Essential for a proper risk assessment.
  $\Longrightarrow$ Study of some important multivariate models such as multivariate normal, spherical and elliptical distributions.
  $\Longrightarrow$ Study of copulas, dependence concepts and dependence measures (including asymptotic dependence vs asymptotic independence).
- Distribution of the sample covariance matrix and the Wishart distribution. Related T^2 Hotelling distribution.
- Multivariate hypothesis testing and simultaneous confidence intervals.
- Analysis of dependence between several variables in a multivariate regression framework. E.g., useful to have some idea of how grades at some exams ("dependent" or "response" variables) are linked to grades at other exams or other variables such as age or gender ("explanatory" variables).

Multivariate problems and methods

- Useful to consider linear combinations of the different variables (e.g., for dimension reduction). One searches for linear combinations which are optimal in some sense (depending on the purpose).
  - If all variables fall in one group, then principal component analysis (PCA) and factor analysis (FA) can help to answer such questions.
  - If the variables fall into more than one group (e.g, “open-book” and “closed-book”), canonical correlation analysis (CCA) is a common method. It uses linear combinations within each group separately while considering the relationship between the two groups of variables.

Multivariate problems and methods

- Imagine that there are g established groups of individuals and that we want to allocate a new individual to one of these groups based on the observations of its characteristics. This is the purpose of discriminant analysis (DA).
  E.g., consider three (50×4) data matrices, each of which concerns measurements on 50 irises of species *Iris setosa*, *Iris versicolour*, and *Iris virginica*, respectively. The variables are

$$X_1 = \text{sepal length}, \quad X_2 = \text{sepal width},$$

$$X_3 = \text{petal length}, \quad X_4 = \text{petal width}.$$

  If a new iris of unknown species has measurements $x_1 = 5.1$, $x_2 = 3.2$, $x_3 = 2.7$ and $x_4 = 0.7$, to which species does it belong?

- Imagine that we have to establish the groups $\Longrightarrow$ cluster analysis (CA).

- Basics about graphical models and directed acyclic graphs.

Two types of methods

Roughly speaking, one can classify the traditional methods of multivariate analysis in two types:

- Methods that consider a subset of variables as response variables to be explained by the remaining variables considered as explanatory variables (e.g., multivariate regression and CCA).
- Methods that analyse the whole set of variables without any distinction (PCA, FA, CA).

In machine learning, this distinction corresponds to the distinction between supervised and unsupervised learning.

Approach

Some references

- Mardia, K.V., Kent, J. T., and Bibby, J. M. (1979), *Multivariate Analysis*, Academic Press.
- Anderson, T. W. (2003), *An Introduction to Multivariate Statistical Analysis*, Wiley.
- McNeil, A. J., Frey, R. and Embrechts, P. (2015), *Quantitative Risk Management*, Princeton University Press.
- Flury, B. (1997), *A First Course in Multivariate Statistics*, Springer.
- Muirhead, R. J. (2005), *Aspects of Multivariate Statistical Theory*, Wiley.

- Course mainly based on the three first references above.
- Completely renewed version of the course:
  - On top of traditional topics of multivariate analysis, deep focus on multivariate models and copulas, and, if possible, basics of graphical models.
  - Connection with real-world problems, especially in quantitative risk management. Applications of the methods to concrete data.
- Exercises: theory, programming (implementation of some methods), applications of the methods to real data examples.

Outline

Distribution of a random vector

- Let $\mathbf{X} = (X_1, \ldots, X_p)' : \Omega \to \mathbb{R}^p$ be a random vector.
- Its joint distribution function (df) is, for $x_1, \ldots, x_p \in \mathbb{R}^p$,

$$F_{\mathbf{X}}(x_1, \ldots, x_p) = \mathbb{P}(X_1 \leq x_1, \ldots, X_p \leq x_p) = \mathbb{P}(\mathbf{X} \leq \mathbf{x}).$$

- $F_{\mathbf{X}}$ denoted F if no ambiguity.
- The df F is said to be absolutely continuous if, for any $x_1, \ldots, x_p \in \mathbb{R}$,

$$F(x_1, \ldots, x_p) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_p} f(u_1, \ldots, u_p) \mathrm{d}u_1 \ldots \mathrm{d}u_p,$$

  for some non-negative function f which is termed the joint probability density function (pdf) of $\mathbf{X}$.
- We have

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(u_1, \ldots, u_p) \mathrm{d}u_1 \ldots \mathrm{d}u_p = 1.$$

- For any $D \subseteq \mathbb{R}^p$,

$$\mathbb{P}(\mathbf{X} \in D) = \int_D f(\mathbf{u}) \mathrm{d}\mathbf{u}.$$

- At any point where F is differentiable, we have

$$f(x_1, \ldots, x_p) = \frac{\partial^p}{\partial x_1 \ldots \partial x_p} F(x_1, \ldots, x_p).$$

Marginal distributions

- The marginal df of X_i, $i = 1, \ldots, p$, denoted by F_{X_i} or F_i, satisfies

$$F_i(x_i) = \mathbb{P}(X_i \leq x_i) = F(\infty, \ldots, \infty, x_i, \infty, \ldots, \infty).$$

- Similarly for k-dimensional margins. Suppose we partition $\mathbf{X}$ into $(\mathbf{X}_1', \mathbf{X}_2')'$, where $\mathbf{X}_1 = (X_1, \ldots, X_k)'$ and $\mathbf{X}_2 = (X_{k+1}, \ldots, X_p)'$, then the marginal df of $\mathbf{X}_1$ is

$$F_{\mathbf{X}_1}(\mathbf{x}_1) = \mathbb{P}(\mathbf{X}_1 \leq \mathbf{x}_1) = F(x_1, \ldots, x_k, \infty, \ldots, \infty).$$

- The marginal pdf of X_i, denoted by f_{X_i} or f_i is given by

$$f_i(x_i) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(u_1, \ldots, u_{i-1}, x_i, u_{i+1}, \ldots, u_p) \mathrm{d}u_1 \ldots \mathrm{d}u_{i-1} \mathrm{d}u_{i+1} \ldots \mathrm{d}u_p,$$

  $x_i \in \mathbb{R}$. Easy extension to the case of k-dimensional pdfs, $k = 2, \ldots, p-1$.
- Existence of a joint pdf $\implies$ Existence of marginal pdfs for all k-dimensional marginals, with $k = 1, \ldots, p-1$. But converse false in general.

Survival function

- Sometimes useful to work with the survival function of $\mathbf{X}$, defined by

$$\bar{F}_{\mathbf{X}}(\mathbf{x}) = \mathbb{P}(\mathbf{X} \geq \mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^p.$$

- Denoted $\bar{F}$ if no ambiguity.
- The marginal survival function of X_i, $i = 1, \ldots, p$, written $\bar{F}_{X_i}$ or simply $\bar{F}_i$ is given by

$$\bar{F}_i(x_i) = \mathbb{P}(X_i > x_i) = \bar{F}(-\infty, \ldots, -\infty, x_i, -\infty, \ldots, -\infty), \quad x_i \in \mathbb{R}.$$

- Note that $\bar{F}(\mathbf{x}) \neq 1 - F(\mathbf{x})$ in general (unless $p = 1$).
- By replacing integrals by sums, one obtains similar formulas for the discrete case, in which the notion of densities is replaced by probability mass functions.

Conditional distributions and independence

- A multivariate model in the form of a joint df, pdf or survival function, implicitly describes the dependence of $X_1, \ldots, X_p \implies$ We can make statements about conditional probabilities and independence.
- Let $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$ as above and assume absolute continuity of the df of $\mathbf{X}$. The conditional distribution of $\mathbf{X}_2$ given $\mathbf{X}_1 = \mathbf{x}_1$ has pdf

$$f_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) = \frac{f(\mathbf{x}_1, \mathbf{x}_2)}{f_{\mathbf{X}_1}(\mathbf{x}_1)},$$

  and the corresponding df is

$$F_{\mathbf{X}_2|\mathbf{X}_1}(\mathbf{x}_2|\mathbf{x}_1) = \int_{u_{k+1}=-\infty}^{x_{k+1}} \cdots \int_{u_p=-\infty}^{x_p} \frac{f(x_1, \ldots, x_k, u_{k+1}, \ldots, u_p)}{f_{\mathbf{X}_1}(\mathbf{x}_1)} \mathrm{d}u_{k+1} \ldots \mathrm{d}u_p.$$

- $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent if and only if (iff) for any $\mathbf{x} = (\mathbf{x}_1', \mathbf{x}_2')'$,

$$F(\mathbf{x}) = F_{\mathbf{X}_1}(\mathbf{x}_1) F_{\mathbf{X}_2}(\mathbf{x}_2),$$

  or, if $\mathbf{X}$ possesses a pdf,

$$f(\mathbf{x}) = f_{\mathbf{X}_1}(\mathbf{x}_1) f_{\mathbf{X}_2}(\mathbf{x}_2).$$

- The individual components of $\mathbf{X}$ are mutually independent iff $F(\mathbf{x}) = \prod_{i=1}^p F_i(x_i)$ for all $\mathbf{x} \in \mathbb{R}^p$, or if $\mathbf{X}$ has a density, $f(\mathbf{x}) = \prod_{i=1}^p f_i(x_i)$.

Moments

- If $\mathbb{E}[|X_i|] < \infty$, $i = 1, \ldots, p$, the mean vector of $\mathbf{X}$ is defined by

$$\boldsymbol{\mu} = \mathrm{E}(\mathbf{X}) = (\mathbb{E}(X_1), \ldots, \mathbb{E}(X_p))'.$$

- If $\mathbb{E}[X_i^2]$, $i = 1, \ldots, p$, the covariance matrix of $\mathbf{X}$ is

$$\Sigma = \mathrm{V}(\mathbf{X}) = \mathrm{E}[(\mathbf{X} - \mathrm{E}(\mathbf{X}))(\mathbf{X} - \mathrm{E}(\mathbf{X}))'] = \mathrm{E}(\mathbf{X}\mathbf{X}') - \mathrm{E}(\mathbf{X})\mathrm{E}(\mathbf{X})'.$$

  Its (i, j)-th element is $\mathrm{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$.

- Introducing the standardized vector $\mathbf{Y}$ with components $Y_i = X_i / \sqrt{\mathrm{Var}(X_i)}$, $i = 1, \ldots, p$, one can define the correlation matrix of $\mathbf{X}$, $\rho(\mathbf{X}) = \mathrm{V}(\mathbf{Y})$.

- The cross covariance matrix between a p-dimensional rv $\mathbf{X}$ and a q-dimensional rv $\mathbf{Y}$ is defined by the $(p \times q)$ matrix

$$\mathrm{C}(\mathbf{X}, \mathbf{Y}) = \mathrm{E}[(\mathbf{X} - \mathrm{E}(\mathbf{X}))(\mathbf{Y} - \mathrm{E}(\mathbf{Y}))'].$$

  Obviously, $\mathrm{C}(\mathbf{X}, \mathbf{X}) = \mathrm{V}(\mathbf{X})$.

Some notations

- Capital letters in normal font for random variables.
- Capital letters in bold font for random vectors.
- Small letters for their realizations.
- Capital in normal font for random matrices AND their realizations.
- Bold and normal font for deterministic vectors and matrices, respectively.
- $\mathbb{E}$, Var, Cov and Corr denote expectation, variance, covariance and correlation in the case of random variables.
- E, V and C denote mean vector, covariance matrix and cross covariance matrix (case of random vectors).
- I_p denotes the identity matrix of dimension $p \times p$.
- $0_{n \times p}$ denotes the zero matrix of dimension $n \times p$.
- 0_p denotes the zero matrix of dimension $p \times p$.
- $\mathbf{1}_n$ denotes the n-dimensional vector $(1, \ldots, 1)'$.

Some properties of E, V and C

- For any $A \in \mathbb{R}^{n\times p}$ and $\mathbf{b} \in \mathbb{R}^n$,

$$\mathrm{E}(A\mathbf{X} + \mathbf{b}) = A\mathrm{E}(\mathbf{X}) + \mathbf{b}$$

and

$$\mathrm{V}(A\mathbf{X} + \mathbf{b}) = A\mathrm{V}(\mathbf{X})A'.$$

- For any $\mathbf{a} \in \mathbb{R}^p$,

$$0 \leq \mathrm{Var}(\mathbf{a}'\mathbf{X}) = \mathbf{a}'\mathrm{V}(\mathbf{X})\mathbf{a},$$

yielding that covariance matrices are positive semidefinite.

- If $\mathbf{a}'\mathrm{V}(\mathbf{X})\mathbf{a} > 0$ for any $\mathbf{a} \in \mathbb{R}^p \backslash \{\mathbf{0}\}$, $\mathrm{V}(\mathbf{X})$ is positive definite and can be shown to be invertible.
- For p and q-dimensional rvs $\mathbf{X}, \mathbf{Y}$ and $A \in \mathbb{R}^{n\times p}$, $B \in \mathbb{R}^{n\times q}$,

$$\mathrm{C}(\mathbf{X}, \mathbf{Y}) = 0_{p\times q} \quad \text{if} \quad \mathbf{X}, \mathbf{Y} \text{ independent (converse not true)},$$

$$\mathrm{C}(\mathbf{X}, \mathbf{Y}) = \mathrm{C}(\mathbf{Y}, \mathbf{X})',$$

$$\mathrm{C}(A\mathbf{X}, B\mathbf{Y}) = A\mathrm{C}(\mathbf{X}, \mathbf{Y})B',$$

and, if $p = q$,

$$\mathrm{E}(\mathbf{X} + \mathbf{Y}) = \mathrm{E}(\mathbf{X}) + \mathrm{E}(\mathbf{Y}),$$

$$\mathrm{V}(\mathbf{X} + \mathbf{Y}) = \mathrm{V}(\mathbf{X}) + \mathrm{V}(\mathbf{Y}) + \mathrm{C}(\mathbf{X}, \mathbf{Y}) + \mathrm{C}(\mathbf{Y}, \mathbf{X}).$$

Standard estimators of covariance and correlation

- Assume that $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are identically distributed from a distribution with mean vector $\boldsymbol{\mu}$, covariance matrix Σ and correlation matrix P.
- Standard estimators of $\boldsymbol{\mu}$, Σ and P are

$$\bar{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i \quad \text{(sample mean vector)}.$$

$$S = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})' \quad \text{(sample covariance matrix)}.$$

$R = (R_{i,j})$, where $R_{i,j} = \dfrac{S_{ij}}{\sqrt{S_{ii}S_{jj}}}$ (sample correlation matrix).

- $\bar{\mathbf{X}}$ is unbiased but S is biased. An unbiased version is $S_n = nS/(n-1)$.
- Other properties depend on the true multivariate distribution of the observations.
- If $\mathbf{X}_1, \ldots, \mathbf{X}_n$ are iid multivariate normal (see later) with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ, then $\bar{\mathbf{X}}$ and S are the maximum likelihood estimators (MLEs) of $\boldsymbol{\mu}$ and Σ.

Higher and conditional moments

Higher moments

Let $\mathbf{X}$ be a p-dimensional rv with mean vector μ and covariance matrix Σ. The multivariate skewness is defined by

$$\beta_{1,p} = \mathbb{E}[\{(\mathbf{X} - \mu)'\Sigma^{-1}(\mathbf{Y} - \mu)\}^3],$$

and the multivariate kurtosis is

$$\beta_{2,p} = \mathbb{E}[\{(\mathbf{X} - \mu)'\Sigma^{-1}(\mathbf{Y} - \mu)\}^2],$$

where $\mathbf{X}$ and $\mathbf{Y}$ are independent and identically distributed.

These measures are invariant under linear transformations.

Conditional moments

For two rvs $\mathbf{X}_1, \mathbf{X}_2$, moments of $\mathbf{X}_1|\mathbf{X}_2$ are called conditional moments. In particular, $\mathrm{E}(\mathbf{X}_1|\mathbf{X}_2)$ and $\mathrm{V}(\mathbf{X}_1|\mathbf{X}_2)$ are the conditional mean vector and the conditional covariance matrix of $\mathbf{X}_1|\mathbf{X}_2$.

Characteristic function

Definition: characteristic function

Let $\mathbf{X}$ be a p-dimensional random vector. Its characteristic function (cf) is defined by

$$\phi_{\mathbf{X}}(\mathbf{t}) = \mathbb{E}[\exp(i\mathbf{t}'\mathbf{X})], \quad \mathbf{t} \in \mathbb{R}^p.$$

- Very useful as it completely characterizes the distribution of a rv.
- Hence many properties of rvs can be shown using the characteristic function.
- Reminder: if a univariate random variable X follows the normal distribution with mean μ and variance σ^2, then

$$\phi_X(t) = \exp\left(it\mu - \frac{1}{2}\sigma^2 t^2\right), \quad t \in \mathbb{R}.$$

Characteristic function

Properties

- The cf always exists, $\phi_{\mathbf{X}}(\mathbf{0}) = 1$ and $|\phi_{\mathbf{X}}(\mathbf{t})| \leq 1$.
- Two rvs have the same cf iff they have the same distribution (uniqueness theorem).
- If the cf $\phi_{\mathbf{X}}(\mathbf{t})$ is absolutely integrable, then $\mathbf{X}$ has pdf

$$f(\mathbf{x}) = \frac{1}{(2\pi)^p} \int_{\mathbb{R}^p} \exp(-i\mathbf{t}'\mathbf{x})\phi_{\mathbf{X}}(\mathbf{t})\mathrm{d}\mathbf{t};$$

  (inversion theorem).
- Let $\mathbf{X} = (\mathbf{X}_1', \mathbf{X}_2')'$. The rvs $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent iff

$$\phi_{\mathbf{X}}(\mathbf{t}) = \phi_{\mathbf{X}_1}(\mathbf{t}_1)\phi_{\mathbf{X}_2}(\mathbf{t}_2),$$

  where $\mathbf{t} = (\mathbf{t}_1', \mathbf{t}_2')'$.
- If $\mathbf{X} = (X_1, \ldots, X_p)'$, then

$$X_1, \ldots, X_p \text{ mutually independent } \Leftrightarrow \phi_{\mathbf{X}}(\mathbf{t}) = \prod_{i=1}^{p} \phi_{X_i}(t_i) \ \forall \mathbf{t} = (t_1, \ldots, t_p)'.$$

- If $\mathbf{X}$ and $\mathbf{Y}$ are independent p-dimensional rvs, then

$$\phi_{\mathbf{X}+\mathbf{Y}}(\mathbf{t}) = \phi_{\mathbf{X}}(\mathbf{t})\phi_{\mathbf{Y}}(\mathbf{t}).$$

Transformations

Theorem

Let $\mathbf{X}$ be a p-dimensional rv with pdf $f_{\mathbf{X}}(\mathbf{x})$ and let $\mathbf{X} = u(\mathbf{Y})$ be a transformation from a k-dimensional rv $\mathbf{Y}$ to $\mathbf{X}$ which is one-to-one except possibly on sets of Lebesgue measure 0 in the supports of $\mathbf{X}$ and $\mathbf{Y}$. Let J be the Jacobian of the transformation from $\mathbf{Y}$ to $\mathbf{X}$, i.e.,

$$J = \det\left(\left(\frac{\partial}{\partial y_j} u_i(y_1, \ldots, y_k)\right)_{i,j}\right).$$

Then the pdf of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(u(\mathbf{y}))|J|.$$

Example: linear transformation

Let $\mathbf{Y} = A\mathbf{X} + \mathbf{b}$, where A is a non-singular matrix (i.e, with non-zero determinant). We have $\mathbf{X} = A^{-1}(\mathbf{Y} - \mathbf{b})$ and therefore $\partial x_i / \partial y_j = (A^{-1})_{ij}$. Thus,

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(A^{-1}(\mathbf{y} - \mathbf{b}))|\det(A^{-1})|.$$

Cramér-Wold device

Theorem (Cramér-Wold)

The distribution of a p-dimensional random vector $\mathbf{X}$ is completely determined by the set of all one-dimensional distributions of linear combinations $\mathbf{t}'\mathbf{X}$, $\mathbf{t} \in \mathbb{R}^p$.

Proof.

For $\mathbf{t} \in \mathbb{R}^p$, let $Y = \mathbf{t}'\mathbf{X}$. The cf of Y is

$$\phi_Y(s) = \mathbb{E}[\exp(isY)] = \mathbb{E}[\exp(is\mathbf{t}'\mathbf{X})].$$

Thus, for $s = 1$,

$$\phi_Y(1) = \mathbb{E}[\exp(iY)] = \mathbb{E}[\exp(i\mathbf{t}'\mathbf{X})],$$

which is the cf of $\mathbf{X}$ at that particular $\mathbf{t}$. □