Some basics on Gaussian measures and variables

The content of this note may be found in any standard textbook on probability theory or statistics.

Table of contents

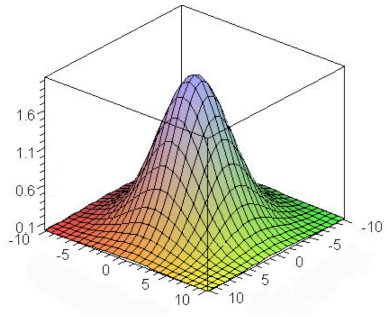
1 Standard Gaussian measure in $\mathbb{R}^n$

The standard Gaussian measure, or normal distribution, γ_n on the Borel sets of $\mathbb{R}^n$ is given by

$$d\gamma_n(x) \,=\, \frac{1}{(2\pi)^{\frac{n}{2}}}\, e^{-\frac{1}{2}|x|^2} d\lambda_n(x)$$

Drafted by M. L. **v1** November 2023

where λ_n is the Lebesgue measure on $\mathbb{R}^n$ (and, for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $|x|^2 = \langle x, x \rangle = \sum_{k=1}^n x_k^2$).



Density of the standard normal distribution in dimension 2

The Gaussian measure γ_n has $\mathbb{R}^n$ as full support, and is equivalent to the Lebesgue measure λ_n. It is not translation invariant: shifted measures are described by

$$\gamma_n(B + h) \;=\; e^{-\frac{1}{2}|h|^2} \int_B e^{-\langle h, x \rangle} d\gamma_n \tag{1}$$

where, for the shift $h \in \mathbb{R}^n$, $B + h = \{x + h; x \in B\}$, B Borel set in $\mathbb{R}^n$. In particular, the shifted measure $\gamma_n(\cdot + h)$ is absolutely continuous (equivalent) with respect to γ_n, with density $e^{-\frac{1}{2}|h|^2 - \langle h, x \rangle}$, $x \in \mathbb{R}^n$.

The Gaussian measure γ_n is the product measure $\gamma_n = \gamma_1 \otimes \cdots \otimes \gamma_1$ of n copies of the one-dimensional measure. The measure γ_n is symmetric (with respect to the origin in $\mathbb{R}^n$), invariant under the action of the orthogonal group, and its push-forward by the map $x \mapsto \frac{x}{|x|}$ is the uniform measure on the unit sphere of $\mathbb{R}^n$. It is log-concave (the logarithm of the density is concave, actually its derivative is uniformly bounded from above by – equal to – -1).

It is already of interest to emphasize at this stage a basic integration by parts formula

$$\int_{\mathbb{R}^n} x f \, d\gamma_n \;=\; \int_{\mathbb{R}^n} \nabla f \, d\gamma_n \tag{2}$$

(as vector integrals) for any smooth enough function $f : \mathbb{R}^n \to \mathbb{R}$ (locally Lipschitz, such that $|x||f|$ and $|\nabla f|$ are integrable). This immediately follows, in dimension one, from the fact that $e^{-\frac{1}{2}x^2}$ is the anti-derivative of $-x\, e^{-\frac{1}{2}x^2}$.

The distribution function of γ_1,

$$\Phi(t) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}x^2} d\lambda_1(x), \quad t \in \mathbb{R}, \tag{3}$$

is continuous, strictly increasing. By symmetry, $1-\Phi(t)=\Phi(-t)$ for every $t\geq 0$ (in particular $\Phi(0)=\frac{1}{2}$). It is convenient to set $\Phi(-\infty)=0$, $\Phi(+\infty)=1$. In the same way, the well-defined inverse function $\Phi^{-1}:[0,1]\to\mathbb{R}$ satisfies $\Phi^{-1}(0)=-\infty$ and $\Phi^{-1}(1)=+\infty$. A basic (although not the sharpest) two-sided bound on Φ is given by

$$\Big(\frac{1}{t}-\frac{1}{t^3}\Big)\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}t^2}\;\leq\;1-\Phi(t)\;\leq\;\frac{1}{t}\,\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}t^2} \tag{4}$$

for every $t>0$ (use integration by parts on $\int_t^\infty \frac{1}{x^k}\,e^{-\frac{1}{2}x^2}d\lambda_1(x)$, $k=2,4$).

The upper-bound (by a simple study of function)

$$1-\Phi(t)\;\leq\;\frac{1}{2}\,e^{-\frac{1}{2}t^2},\quad t\geq 0, \tag{5}$$

is already quite useful.

The associated *error function* is $\mathcal{E}(t)=\frac{1}{\sqrt{2\pi}}\int_{-t}^{t}e^{-\frac{1}{2}x^2}d\lambda_1(x)=2\Phi(t)-1$, $t\geq 0$. By the preceding, $1-\mathcal{E}(t)\leq e^{-\frac{1}{2}t^2}$.

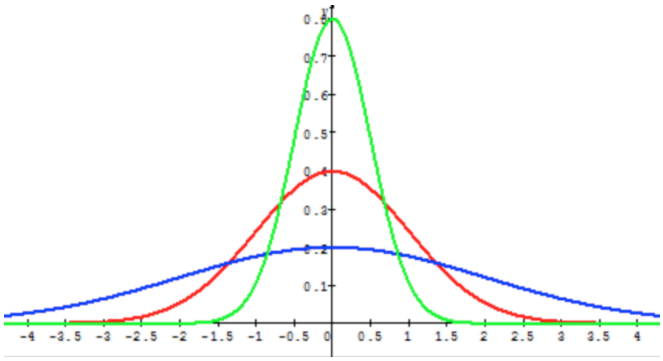
The Gaussian measure γ_n generates all other Gaussian measures, in finite as well as infinite, dimension. This is best explained via random variables and vectors.

2 Real Gaussian random variable

A real random variable X on a probability space $(\Omega,\mathcal{A},\mathbb{P})$ is said to have a Gaussian, or normal, distribution $\mathcal{N}(m,\sigma^2)$ with parameters $m\in\mathbb{R}$ and $\sigma^2>0$, if its law has density

$$f_{m,\sigma^2}(x)\;=\;\frac{1}{\sqrt{2\pi\sigma^2}}\,e^{-\frac{1}{2\sigma^2}(x-m)^2},\quad x\in\mathbb{R},$$

with respect to the Lebesgue measure on $\mathbb{R}$. The terminology of the parameters is due to the fact that the mean of the law of X is $\mathbb{E}(X)=m$ and its variance $\mathrm{Var}(X)=\sigma^2$, which thus completely determine the distribution within the class of Gaussian random variables.



It is sometimes convenient to include the degenerate case $\sigma = 0$ into the picture, corresponding to the Dirac mass at m.

The standard normal distribution $\mathcal{N}(0,1)$, that is γ_1 in the notation of the first section, generates all the other normal distributions, since if Y has law $\mathcal{N}(0,1)$, then $X = m + \sigma Y$ has law $\mathcal{N}(m, \sigma^2)$ (and conversely, if X has law $\mathcal{N}(m, \sigma^2)$, then $Y = \frac{X-m}{\sigma}$ has law $\mathcal{N}(0,1)$). This observation follows from an affine change of variable in the definition of the Gaussian density.

The Laplace transform of a random variable Y with distribution $\mathcal{N}(0,1)$ is easily computed as

$$\mathbb{E}\big(e^{uY}\big) \;=\; \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ux - \frac{1}{2}x^2} d\lambda_1(x) \;=\; e^{\frac{1}{2}u^2}, \quad u \in \mathbb{R}.$$

As a consequence, if X has law $\mathcal{N}(m, \sigma^2)$, then $\mathbb{E}(e^{uX}) \;=\; e^{mu + \frac{1}{2}\sigma^2 u^2}$, $u \in \mathbb{R}$. A real Gaussian random variable thus admits moments of all orders. More precisely, a Taylor expansion on the Laplace transform indicates that if, for example, Y has law $\mathcal{N}(0,1)$,

$$\mathbb{E}\big(e^{uY}\big) \;=\; \sum_{k=0}^{\infty} \frac{u^k}{k!}\, \mathbb{E}(Y^k).$$

Identification with the series defining $e^{\frac{1}{2}u^2}$ yields

$$\mathbb{E}(Y^{2k}) \;=\; \frac{(2k)!}{2^k k!} \tag{6}$$

and $\mathbb{E}(Y^{2k+1}) = 0$, $k \in \mathbb{N}$. These equations also follow, recursively, from the integration by parts formula (2) $\mathbb{E}(Yf(Y)) = \mathbb{E}(f'(Y))$ applied to $f(x) = x^{2k+1}$.

As a random variable X with law $\mathcal{N}(0, \sigma^2)$ may be represented by σY where Y is distributed according to $\mathcal{N}(0,1)$, it follows that all the moments $\mathbb{E}(|X|^p)$, $p > 0$, of X are proportional to σ^p, and thus all their $\frac{1}{p}$-powers are proportional.

The cumulants are the coefficients in the series expansion of the logarithm of the Laplace transform $\log \mathbb{E}(^{uX})$, $u \in \mathbb{R}$. Clearly, for X with distribution $\mathcal{N}(m, \sigma^2)$, the first two coefficients are m and σ^2, and all the others vanish.

It is possible to work equivalently with the Fourier transform, or characteristic function,

$$\varphi_Y(u) \;=\; \mathbb{E}(e^{iuY}) \;=\; e^{-\frac{1}{2}u^2}, \quad u \in \mathbb{R}.$$

(If X has law $\mathcal{N}(m, \sigma^2)$, then $\varphi_X(u) = \mathbb{E}(e^{iuX}) \;=\; e^{imu - \frac{1}{2}\sigma^2 u^2}$, $u \in \mathbb{R}$.) The expression of φ_Y may for example be deduced from the differential equation $\varphi_Y'(u) = -u\,\varphi_Y(u)$, $u \in \mathbb{R}$, as a consequence of (2). The Gaussian density is therefore the fixed point of the Fourier transform.

From either a direct convolution argument on the Gaussian density, or the Fourier transform, if X_1 et X_2 are independent with respective normal distributions $\mathcal{N}(m_1, \sigma_1^2)$ and $\mathcal{N}(m_2, \sigma_2^2)$, $m_1, m_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, then $X_1 + X_2$ has law $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$. In particular, if $X_1, \ldots, X_n$ are independent random variables with common normal law $\mathcal{N}(m, \sigma^2)$, then

$$\frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (X_k - m)$$

has law $\mathcal{N}(0, 1)$. Cramér's theorem shows conversely that if X_1 and X_2 are two independent random variables variables with $X_1 + X_2$ normally distributed, then both X_1 and X_2 follow a normal distribution.

A simple check on the variances indicates that if X and Y are independent centered Gaussian variables, for any $\theta \in \mathbb{R}$, $X(\theta) = X\sin(\theta) + Y\cos(\theta)$ and $X'(\theta) = X\cos(\theta) - Y\sin(\theta)$ are Gaussian, independent with the same law as X. That is, the couples (X, Y) and $(X(\theta), X'(\theta))$ have the same distribution. Bernstein's theorem ensures conversely that if X and Y are independent and $X+Y$ and $X-Y$ are also independent, then both X and Y must have normal distributions. The statement actually holds true for the couple $(X(\theta), X'(\theta))$ provided that θ is not an integer multiple of $\frac{\pi}{2}$.

3 Gaussian random vector

A random vector $X = (X_1, \ldots, X_n)$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with values in $\mathbb{R}^n$ is said to be Gaussian if every linear combination of its coordinates is a real Gaussian random variable. That is, for every $c = (c_1, \ldots, c_n) \in \mathbb{R}^n$,

$$\langle c, X \rangle \;=\; \sum_{k=1}^{n} c_k X_k$$

is a real Gaussian variable.

More generally, a sequence $(X_k)_{k\geq 1}$ of (real) random variables is said to be Gaussian if every finite subsequence is a Gaussian vector.

Since Gaussian, the distribution of $\langle c, X \rangle$ is determined by its mean $\mathbb{E}(\langle c, X \rangle)$ and variance $\mathrm{Var}(\langle c, X \rangle)$. By linearity,

$$\mathbb{E}\big(\langle c, X \rangle\big) \;=\; \sum_{k=1}^{n} c_k\, \mathbb{E}(X_k) \;=\; \langle c,\, \mathbb{E}(X) \rangle$$

where $m = \mathbb{E}(X) = (\mathbb{E}(X_1), \ldots, \mathbb{E}(X_n))$ is the vector of the means, and by bilinearity,

$$\mathrm{Var}\big(\langle c, X \rangle\big) \;=\; \sum_{k,\ell=1}^{n} c_k c_\ell\, \mathbb{E}\big(\big[X_k - \mathbb{E}(X_k)\big]\big[X_\ell - \mathbb{E}(X_\ell)\big]\big) \;=\; \langle \Sigma c, c \rangle$$

where $\Sigma = (\Sigma_{k,\ell})_{1\leq k,\ell\leq n}$,

$$\Sigma_{k,\ell} \,=\, \mathbb{E}\big(\big[X_k - \mathbb{E}(X_k)\big]\big[X_\ell - \mathbb{E}(X_\ell)\big]\big),$$

is the covariance matrix of the law of X. Hence, the law of the Gaussian random vector X is described by the mean vector $m = \mathbb{E}(X)$ and the covariance matrix Σ, and usually denoted, by analogy with the real case, by $\mathcal{N}(m, \Sigma)$.

In terms of the Fourier transform of the law of X,

$$\varphi_X(u) \,=\, \mathbb{E}\big(e^{i\langle u,X\rangle}\big) \,=\, \varphi_{\langle u,X\rangle}(1) \,=\, e^{i\langle m,u\rangle - \frac{1}{2}\langle \Sigma u,u\rangle}, \quad u \in \mathbb{R}^n.$$

Any affine transformation of a Gaussian vector is still a Gaussian vector. By the very definition of a Gaussian vector $X = (X_1, \ldots, X_n)$, the coordinates $X_1, \ldots, X_n$ are real Gaussian random variables. But conversely, Gaussian entries do not always ensure that the vector is Gaussian. For example, if Y is a standard normal real random variable on $(\Omega, \mathcal{A}, \mathbb{P})$, and ε an independent Bernoulli random variable with law $\mathbb{P}(\varepsilon = +1) = \mathbb{P}(\varepsilon = -1) = \frac{1}{2}$, then both Y and εY are Gaussian, but the vector $(Y, \varepsilon Y)$ is not Gaussian.

A check on the covariances indicates, as in the real case, that if X and Y are independent centered Gaussian vectors, for any $\theta \in \mathbb{R}$, $X(\theta) = X\sin(\theta) + Y\cos(\theta)$ and $X'(\theta) = X\cos(\theta) - Y\sin(\theta)$ are Gaussian, independent with the same law as X. That is, the couples (X, Y) and $(X(\theta), X'(\theta))$ have the same distribution.

As in the real case, the Gaussian vector with mean zero and covariance matrix $\Sigma = \mathrm{Id}$, the $n \times n$ identity matrix, plays a central role and generates all other Gaussian vectors. Its law is actually the standard Gaussian measure γ_n as described in the first section. Let indeed $Y = (Y_1, \ldots, Y_n)$ be a random vector with law γ_n. The marginal distributions are all the standard one-dimensional Gaussian distribution $\mathcal{N}(0, 1)$, and the density f being a product of the marginals, the random variables $Y_1, \ldots, Y_n$ are independent. In the preceding notation, Y has law $\mathcal{N}(0, \mathrm{Id})$.

A covariance matrix Σ is symmetric and (semi-) positive definite, so that it admits a square root $\Sigma = A\,{}^{\top}\!A$ (with ${}^{\top}\!A$ the transpose of A). Therefore, if Y follows the standard Gaussian law $\mathcal{N}(0, \mathrm{Id})$ in $\mathbb{R}^n$, any Gaussian random vector X in $\mathbb{R}^n$ with mean m and covariance matrix $\Sigma = A\,{}^{\top}\!A$ has the same distribution as $m + AY$.

As a consequence, if Σ, and hence its square root A, are invertible, for any Borel set B in $\mathbb{R}^n$,

$$\mathbb{P}(X \in B) \,=\, \mathbb{P}(m + AY \in B) \,=\, \mathbb{P}\big(Y \in A^{-1}(B - m)\big),$$

so that

$$\mathbb{P}(X \in B) \,=\, \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{A^{-1}(B-m)} e^{-\frac{1}{2}|y|^2}\, d\lambda_n(y).$$

After the change of variable $x = m + Ay$ in $\mathbb{R}^n$ with Jacobian $\det(A)$, it follows that

$$\begin{aligned}\mathbb{P}(X \in B) &= \frac{1}{(2\pi)^{\frac{n}{2}}|\det(A)|}\int_B e^{-\frac{1}{2}|A^{-1}(x-m)|^2}\, d\lambda_n(x)\\ &= \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\det(\Sigma)|}}\int_B e^{-\frac{1}{2}\langle\Sigma^{-1}(x-m),x-m\rangle}\, d\lambda_n(x).\end{aligned}$$

This formula expresses that the law of the Gaussian random vector X, with invertible covariance matrix Σ, admits the density

$$\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\det(\Sigma)|}}\, e^{-\frac{1}{2}\langle\Sigma^{-1}(x-m),x-m\rangle}, \quad x \in \mathbb{R}^n,$$

with respect to the Lebesgue measure λ_n on $\mathbb{R}^n$.

The preceding description of the density of a Gaussian vector assumes the non-singularity of the covariance matrix. In case of singularities, densities are described similarly but in subspaces with dimension the rank of the covariance matrix. (See also below.)

The coordinates $(X_1, \ldots, X_n)$ of a Gaussian random vector $X = (X_1, \ldots, X_n)$ are independent if and only if the covariance matrix Σ of X is diagonal (as can be checked easily, for example, on the Fourier transform).

More generally, given a random vector X with law $\mathcal{N}(0, \mathrm{Id})$ in $\mathbb{R}^n$, Cochran's theorem indicates that if F is a vector subspace in $\mathbb{R}^n$ with orthogonal subspace $F^\perp$, and if P_F and $P_{F^\perp}$ denote the matrices of the orthogonal projections on F and $F^\perp$ respectively, then $P_F X$ and $P_{F^\perp} X$ are independent with respective distributions $\mathcal{N}(0, P_F)$ and $\mathcal{N}(0, P_{F^\perp})$. By induction, a similar result holds for any decomposition of $\mathbb{R}^n$ in orthogonal subspaces.

In the following, let X be a Gaussian random vector in $\mathbb{R}^n$, centered for simplicity. Actually, up to translation, every Gaussian variable or vector can be centered, and this is usually mostly assumed for simplicity.

The covariance matrix Σ of the law of X is symmetric and (semi-) positive definite, so may be diagonalized in an orthogonal basis, with non-negative eigenvalues. That is, there exists a orthogonal matrix Q such that $\Sigma = QD^\top Q$, where D is diagonal, with non-negative (eigen-) values $s_1, \ldots, s_n \geq 0$ on the diagonal (some directions might be degenerated). In this representation, $A = Q\sqrt{D}$, where $\sqrt{D}$ is the diagonal matrix consisting of $\sqrt{s_1}, \ldots, \sqrt{s_n}$. Hence the Gaussian vector X has the same distribution as

$$AY \;=\; Q\sqrt{D}\, Y$$

where Y follows the standard normal law $\mathcal{N}(0, \mathrm{Id})$. Now, the (Gaussian) vector $\sqrt{D}\, Y$, with diagonal covariance matrix D, admits independent coordinates. As a consequence of this

construction, starting from the Gaussian vector X, after the change of basis by the matrix ${}^\top Q$, the new Gaussian vector $Z = {}^\top QX$ has the same law as

$$ {}^\top QAY \;=\; {}^\top Q\, Q\sqrt{D}\, Y \;=\; \sqrt{D}\, Y, $$

and its components are therefore independent. Thus, after a suitable change of basis in $\mathbb{R}^n$ built from the covariance matrix, a (centered) Gaussian vector may be put in a position where all its coordinates are independent, which is of course of significant benefit.

Finally, it is useful in applications to benefit from the integration by parts formula (2) for an arbitrary centered Gaussian vector X with covariance matrix $\Sigma = A\,{}^\top A$. Namely, if $f : \mathbb{R}^n \to \mathbb{R}$ is sufficiently smooth,

$$ \mathbb{E}\big(Xf(X)\big) \;=\; A\,\mathbb{E}\big(Yf \circ A(Y)\big) $$

where Y has distribution $\mathcal{N}(0, \mathrm{Id})$, so that

$$ \mathbb{E}\big(Xf(X)\big) \;=\; \Sigma\,\mathbb{E}\big(\nabla f(X)\big) \tag{7} $$

(as vectors in $\mathbb{R}^n$).

4 Convergence of sequences of Gaussian variables

If $(X_k)_{k\in\mathbb{N}}$ is a sequence of Gaussian random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ converging in law to some random variable X, necessarily X is Gaussian. Indeed, if X_k has law $\mathcal{N}(m_k, \sigma_k^2)$, $k \in \mathbb{N}$, it may be shown first that the sequence $(m_k)_{k\in\mathbb{N}}$ is bounded. If not, for any $M > 0$ there exists k such that $|m_k| \geq M$, and in particular

$$ \mathbb{P}\big(|X_k| \geq M\big) \;\geq\; \mathbb{P}\big(|X_k| \geq m_k\big) \;\geq\; \frac{1}{2} $$

since m_k is the median of X_k. But this is incompatible with the tightness of the sequence $(X_k)_{k\in\mathbb{N}}$. By the convergence of the characteristic functions

$$ \lim_{k\to\infty} e^{im_k u - \frac{1}{2}\sigma_k^2 u^2} \;\to\; \varphi_X(u) $$

for every $u \in \mathbb{R}$, where φ_X is the characteristic function of X, so that, by taking modulus, the sequence $(\sigma_k)_{k\in\mathbb{N}}$ is convergent.

There is therefore a subsequence k' of integers such that $m_{k'} \to m \in \mathbb{R}$ and $\sigma_{k'} \to \sigma \in [0, \infty)$, and by the convergence of the characteristic functions, $\varphi_X(u) = e^{imu - \frac{1}{2}\sigma^2 u^2}$ for every $u \in \mathbb{R}$. Hence X has law $\mathcal{N}(m, \sigma^2)$ (possibly degenerated).

In case the sequence $(X_k)_{k\in\mathbb{N}}$ converges almost surely, or only in probability, towards X, it may be shown in addition that $(X_k)_{k\in\mathbb{N}}$ converges to X in any L^p-space, $p > 0$. Namely, from the boundedness of the sequences $(m_k)_{k\in\mathbb{N}}$ and $(\sigma_k)_{k\in\mathbb{N}}$, since all moments of Gaussian random variables are equivalent, $\sup_{k\in\mathbb{N}} \mathbb{E}(|X_k|^p) < \infty$ for every $p > 0$. In particular, the sequence $(|X_k|^p)_{k\in\mathbb{N}}$ is uniformly integrable, from which the claim follows. Of course, this is no more true if the convergence is only in distribution (take for example $X_k = Y$ for all $k \in \mathbb{N}$ where Y is a standard normal variable, and $X = -Y$).

5 Conditioning of Gaussian vectors

Let $(X_1, \ldots, X_n)$ be a Gaussian random vector in $\mathbb{R}^n$, defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, centered for simplicity. For $1 \leq k \leq n-1$, set $U = (X_1, \ldots, X_k)$ and $V = (X_{k+1}, \ldots, X_n)$. The Gaussian structure allows for an easy and complete description of the conditional law of U given $V = v$ $(\in \mathbb{R}^{n-k})$.

Towards this claim, it may first be noticed that, for any $c \in \mathbb{R}^k$, the conditional expectation $\mathbb{E}(\langle c, U\rangle \,|\, V)$ is simply the orthogonal projection, in $\mathrm{L}^2(\mathbb{P})$, on the linear subspace generated by the coordinates of V, and is in particular Gaussian. Indeed, if R represents this orthogonal projection, it is measurable with respect to V and $\langle c, U\rangle = R + S$ where S is orthogonal with respect to the subspace generated by V. But then S is independent from V, and

$$\mathbb{E}\big(\langle c, U\rangle \,|\, V\big) \,=\, \mathbb{E}\big(R \,|\, V\big) + \mathbb{E}\big(S \,|\, V\big) \,=\, R + \mathbb{E}(S) \,=\, R$$

(by centering).

Next, by a Fourier transform argument, the conditional law of U given $V = v$ may be described by the conditional expectations

$$\mathbb{E}\big(e^{i\langle c, U\rangle} \,|\, V\big), \quad c \in \mathbb{R}^k.$$

Fix $c \in \mathbb{R}^k$. By a covariance argument, it is immediate that $\langle c, U\rangle - \mathbb{E}(\langle c, U\rangle \,|\, V)$ and V are orthogonal, and thus, as jointly Gaussian, independent. As a consequence,

$$\begin{aligned}
\mathbb{E}\big(e^{i\langle c, U\rangle} \,|\, V\big) \,&=\, e^{i\,\mathbb{E}(\langle c, U\rangle \,|\, V)}\, \mathbb{E}\Big(e^{i(\langle c, U\rangle - \mathbb{E}(\langle c, U\rangle \,|\, V))} \,|\, V\Big) \\
&=\, e^{i\,\mathbb{E}(\langle c, U\rangle \,|\, V)}\, \mathbb{E}\Big(e^{i(\langle c, U\rangle - \mathbb{E}(\langle c, U\rangle \,|\, V))}\Big) \\
&=\, \exp\Big(i\,\mathbb{E}(\langle c, U\rangle \,|\, V) - \frac{1}{2}\,\mathbb{E}\big(\big[\langle c, U\rangle - \mathbb{E}(\langle c, U\rangle \,|\, V)\big]^2\big)\Big).
\end{aligned}$$

Therefore, by the standard expression of the Fourier transform of Gaussian vectors, the conditional distribution of U given $V = v$ is Gaussian with mean $\mathbb{E}(U \,|\, V = v)$ $(\in \mathbb{R}^k)$ and

covariance matrix (independent on v)

$$\mathbb{E}\Big(\big[U_i - \mathbb{E}(U_i \,|\, V)\big]\big[U_j - \mathbb{E}(U_j \,|\, V)\big]\Big), \quad i, j = 1, \ldots, k.$$

In particular, it is enough to know the conditional expectation $\mathbb{E}(U \,|\, V)$ to determine the conditional laws.

The preceding may also be checked directly on the densities.