On the mean and variance of quadratic functionals of Gaussian random vectors

Alen Alexanderian*

Abstract

We consider quadratic functionals of Gaussian random vectors and derive the expressions for their mean and variance. This note provides a self-contained summary of several well-known facts about the Gaussian distribution, which are then used to derive the desired result.

1 Introduction

We consider the function

$$f(\boldsymbol{x}) = \frac{1}{2}\left\langle \mathbf{F}\boldsymbol{x} + \boldsymbol{c}, \mathbf{Q}(\mathbf{F}\boldsymbol{x} + \boldsymbol{c})\right\rangle, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product, $\boldsymbol{x}$ is an n-dimensional Gaussian random vector, $\boldsymbol{x} \sim N(\boldsymbol{m}, \mathbf{C})$, $\mathbf{F} \in \mathbb{R}^{m\times n}$, $\boldsymbol{c} \in \mathbb{R}^m$, and $\mathbf{Q} \in \mathbb{R}^{m\times m}$ is a symmetric matrix. Our goal is to derive analytic expressions for the mean and variance of f. The derivations rely on results regarding moments of multivariate Gaussian distribution, which have been studied extensively in the probability and statistic literature; see, e.g., [1, 2, 3]. In this note, we require the following:

Proposition 1.1. *Suppose $\boldsymbol{x} \sim N(\mathbf{0}, \mathbf{C})$ is an n-dimensional Gaussian random vector. Then, for $i, j, k, \ell \in \{1, \dots, n\}$,*

(a). $\mathrm{E}\{x_i x_j x_k\} = 0$.
(b). $\mathrm{E}\{x_i x_j x_k x_\ell\} = C_{ij}C_{k\ell} + C_{ik}C_{j\ell} + C_{i\ell}C_{jk}$.

The above result is a special case of general results [1, 2, 3] for moments of the form $\mathrm{E}\{x_{i_1} x_{i_2} \cdots x_{i_s}\}$, where $\{i_1, i_2, \dots . i_s\} \subseteq \{1, \dots, n\}$.

2 The derivations

The following result follows from definition of the covariance matrix and Proposition 1.1.

Proposition 2.1. *Suppose $\boldsymbol{x} \sim N(\mathbf{0}, \mathbf{C})$, and let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix and $\boldsymbol{a}$ and $\boldsymbol{b}$ be fixed vectors in $\mathbb{R}^n$. Then, we have*

(a). $\mathrm{E}\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle \langle \boldsymbol{b}, \boldsymbol{x}\rangle\} = \langle \boldsymbol{a}, \mathbf{C}\boldsymbol{b}\rangle$,
(b). $\mathrm{E}\{\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle\} = \mathrm{tr}(\mathbf{AC})$,
(c). $\mathrm{E}\{\langle \boldsymbol{x}, \boldsymbol{a}\rangle \langle \mathbf{A}\boldsymbol{x}, \boldsymbol{x}\rangle\} = 0$,
(d). $\mathrm{E}\{\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle^2\} = 2\mathrm{tr}\big[(\mathbf{AC})^2\big] + \big[\mathrm{tr}(\mathbf{AC})\big]^2$.

Proof. The statement (a) follows from $\mathrm{E}\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle \langle \boldsymbol{b}, \boldsymbol{x}\rangle\} = \sum_{i,j} a_i b_j \mathrm{E}\{x_i x_j\} = \sum_{i,j} a_i b_j C_{ij} = \langle \boldsymbol{a}, \mathbf{C}\boldsymbol{b}\rangle$. The second statement follows from

$$\mathrm{E}\{\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle\} = \sum_{i,j} A_{ij}\mathrm{E}\{x_i x_j\} = \sum_{i,j} A_{ij}C_{ij} = \sum_i (\mathbf{AC})_{jj} = \mathrm{tr}(\mathbf{AC}).$$

*North Carolina State University, Raleigh, NC, USA. E-mail: alexanderian@ncsu.edu
Last revised: December 20, 2021

For part (c), we note, $\mathrm{E}\{\langle \boldsymbol{x}, \boldsymbol{a}\rangle \langle \mathbf{A}\boldsymbol{x}, \boldsymbol{x}\rangle\} = \sum_{i,j,k} a_i A_{jk} \mathrm{E}\{x_i x_j x_k\} = 0$, where we also used Theorem 1.1(a). To see the statement (d), we note

$$\mathrm{E}\{\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle^2\} = \sum_{i,j,k,\ell=1}^{n} A_{ij} A_{k\ell} \mathrm{E}\{x_i x_j x_k x_\ell\} = \sum_{i,j,k,\ell=1}^{n} A_{ij} A_{k\ell} \Big(C_{ij} C_{k\ell} + C_{ik} C_{j\ell} + C_{i\ell} C_{jk} \Big),$$

where we have used Proposition 1.1(b). We distribute the summation over the three terms in the bracket and note that the first term is given by

$$\sum_{i,j,k,\ell=1}^{n} A_{ij} A_{k\ell} C_{ij} C_{k\ell} = \Big(\sum_{i,j} A_{ij} C_{ij} \Big) \Big(\sum_{k,\ell} A_{k\ell} C_{k\ell} \Big) = \big[\mathrm{tr}(\mathbf{AC})\big]^2. \tag{2.1}$$

Next, we note that the second term $\sum_{i,j,k,\ell=1}^{n} A_{ij} A_{k\ell} C_{ik} C_{j\ell}$ simplifies as follows.

$$\sum_{i,j,k,\ell=1}^{n} A_{ij} A_{k\ell} C_{ik} C_{j\ell} = \sum_{i,k,\ell} \Big(A_{\ell k} C_{ki} \big(\sum_{j} A_{ij} C_{j\ell} \big) \Big) = \sum_{i,\ell} (\mathbf{AC})_{\ell i} (\mathbf{AC})_{i\ell} = \sum_{i} (\mathbf{AC})^2_{ii} = \mathrm{tr}\big[(\mathbf{AC})^2\big].$$

A similar calculation shows that the third term satisfies $\sum_{i,j,k,\ell=1}^{n} A_{ij} A_{k\ell} C_{i\ell} C_{jk} = \mathrm{tr}\big[(\mathbf{AC})^2\big]$, and thus, the result follows. $\square$

Remark 2.2. *Note that parts (a) and (b) of Proposition 2.1 do not require the Gaussian assumption on $\boldsymbol{x}$, and hold for any centered random n-vector with a covariance matrix $\mathbf{C}$.*

It is straightforward to extend Proposition 2.1 to the case of non-centered Gaussian random vectors. This is recorded in the following proposition:

Proposition 2.3. *Suppose $\boldsymbol{x} \sim N(\boldsymbol{m}, \mathbf{C})$, and let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric matrix and $\boldsymbol{a}$ and $\boldsymbol{b}$ be fixed vectors in $\mathbb{R}^n$. We have,*

(a). $\mathrm{E}\{\langle \boldsymbol{a}, \boldsymbol{x}\rangle \langle \boldsymbol{b}, \boldsymbol{x}\rangle\} = \langle \boldsymbol{a}, \mathbf{C}\boldsymbol{b}\rangle + \langle \boldsymbol{a}, \boldsymbol{m}\rangle \langle \boldsymbol{b}, \boldsymbol{m}\rangle$,
(b). $\mathrm{E}\{\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle\} = \mathrm{tr}(\mathbf{AC}) + \langle \boldsymbol{m}, \mathbf{A}\boldsymbol{m}\rangle$,
(c). $\mathrm{E}\{\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle \langle \boldsymbol{x}, \boldsymbol{b}\rangle\} = \big(\mathrm{tr}(\mathbf{AC}) + \langle \boldsymbol{m}, \mathbf{A}\boldsymbol{m}\rangle\big) \langle \boldsymbol{m}, \boldsymbol{b}\rangle + 2\langle \boldsymbol{b}, \mathbf{CA}\boldsymbol{m}\rangle$,
(d). $\mathrm{E}\{\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle^2\} = 2\mathrm{tr}\big[(\mathbf{AC})^2\big] + \mathrm{tr}(\mathbf{AC})^2 + \big[2\mathrm{tr}(\mathbf{AC}) + \langle \boldsymbol{m}, \mathbf{A}\boldsymbol{m}\rangle\big] \langle \boldsymbol{m}, \mathbf{A}\boldsymbol{m}\rangle + 4\langle \mathbf{A}\boldsymbol{m}, \mathbf{CA}\boldsymbol{m}\rangle$.

Proof. Part (a) follows directly from Proposition 2.1(a) and using

$$\langle \boldsymbol{a}, \boldsymbol{x}\rangle \langle \boldsymbol{b}, \boldsymbol{x}\rangle = \langle \boldsymbol{a}, \boldsymbol{x}-\boldsymbol{m}\rangle \langle \boldsymbol{b}, \boldsymbol{x}-\boldsymbol{m}\rangle + \langle \boldsymbol{a}, \boldsymbol{x}\rangle \langle \boldsymbol{b}, \boldsymbol{m}\rangle + \langle \boldsymbol{a}, \boldsymbol{m}\rangle \langle \boldsymbol{b}, \boldsymbol{x}\rangle - \langle \boldsymbol{a}, \boldsymbol{m}\rangle \langle \boldsymbol{b}, \boldsymbol{m}\rangle.$$

Similarly, part (b) follows from $\langle \mathbf{A}\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}-\boldsymbol{m}, \mathbf{A}(\boldsymbol{x}-\boldsymbol{m})\rangle + 2\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{m}\rangle - \langle \boldsymbol{m}, \mathbf{A}\boldsymbol{m}\rangle$ and Proposition 2.1(b). To show part (c), we note that

$$\begin{aligned}&\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle \langle \boldsymbol{x}, \boldsymbol{b}\rangle = \langle \boldsymbol{x}-\boldsymbol{m}, \mathbf{A}(\boldsymbol{x}-\boldsymbol{m})\rangle \langle \boldsymbol{x}-\boldsymbol{m}, \boldsymbol{b}\rangle \\ &+ \langle \boldsymbol{x}, \mathbf{A}\boldsymbol{x}\rangle \langle \boldsymbol{m}, \boldsymbol{b}\rangle + 2\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{m}\rangle \langle \boldsymbol{x}, \boldsymbol{b}\rangle - 2\langle \boldsymbol{x}, \mathbf{A}\boldsymbol{m}\rangle \langle \boldsymbol{m}, \boldsymbol{b}\rangle - \langle \boldsymbol{m}, \mathbf{A}\boldsymbol{m}\rangle \langle \boldsymbol{x}, \boldsymbol{b}\rangle + \langle \boldsymbol{m}, \mathbf{A}\boldsymbol{m}\rangle \langle \boldsymbol{m}, \boldsymbol{b}\rangle,\end{aligned}$$

and use parts (a) and (b) along with Proposition 2.1(c). Proof of part (d), which builds on the results of previous parts and Proposition 2.1, is similar and is ommitted for brevity. $\square$

3 Mean and variance of f

We are ready to derive the expressions for the mean and variance of f defined in (1.1):

Proposition 3.1. *Let $f(\boldsymbol{x})$ be as in (1.1). We have*

(a). $\mathrm{E}\{f(\boldsymbol{x})\} = \frac{1}{2}\mathrm{tr}(\mathbf{HC}) + \frac{1}{2}\langle \boldsymbol{m}, \mathbf{H}\boldsymbol{m}\rangle + \langle \boldsymbol{m}, \boldsymbol{b}\rangle + \dfrac{1}{2}\langle \boldsymbol{c}, \mathbf{Q}\boldsymbol{c}\rangle$,
(b). $\mathrm{var}\{f(\boldsymbol{x})\} = \frac{1}{2}\mathrm{tr}\big[(\mathbf{HC})^2\big] + \langle \mathbf{H}\boldsymbol{m} + \boldsymbol{b}, \mathbf{C}(\mathbf{H}\boldsymbol{m} + \boldsymbol{b})\rangle$,

where $\mathbf{H} = \mathbf{F}^\top \mathbf{QF}$ and $\boldsymbol{b} = \mathbf{F}^\top \mathbf{Q}\boldsymbol{c}$.

Proof. Note that

$$f(\boldsymbol{x}) = \frac{1}{2}\left\langle \mathbf{F}\boldsymbol{x}, \mathbf{QF}\boldsymbol{x}\right\rangle + \left\langle \mathbf{F}\boldsymbol{x}, \mathbf{Q}\boldsymbol{c}\right\rangle + \frac{1}{2}\left\langle \boldsymbol{c}, \mathbf{Q}\boldsymbol{c}\right\rangle = \frac{1}{2}\left\langle \boldsymbol{x}, \mathbf{H}\boldsymbol{x}\right\rangle + \left\langle \boldsymbol{x}, \boldsymbol{b}\right\rangle + \frac{1}{2}\left\langle \boldsymbol{c}, \mathbf{Q}\boldsymbol{c}\right\rangle.$$

Part (a) of the proposition follows from Proposition 2.3. To derive the variance expression, it is convenient to drop the constant term $\frac{1}{2}\left\langle \boldsymbol{c}, \mathbf{Q}\boldsymbol{c}\right\rangle$ and consider $g(\boldsymbol{x}) = \frac{1}{2}\left\langle \boldsymbol{x}, \mathbf{H}\boldsymbol{x}\right\rangle + \left\langle \boldsymbol{x}, \boldsymbol{b}\right\rangle$. We have,

$$\mathrm{E}\{g(\boldsymbol{x})^2\} = \mathrm{E}\Big\{\Big(\frac{1}{2}\left\langle \boldsymbol{x}, \mathbf{H}\boldsymbol{x}\right\rangle + \left\langle \boldsymbol{x}, \boldsymbol{b}\right\rangle\Big)^2\Big\} = \frac{1}{4}\mathrm{E}\{\left\langle \boldsymbol{x}, \mathbf{H}\boldsymbol{x}\right\rangle^2\} + \mathrm{E}\{\left\langle \boldsymbol{x}, \mathbf{H}\boldsymbol{x}\right\rangle \left\langle \boldsymbol{x}, \boldsymbol{b}\right\rangle\} + \mathrm{E}\{\left\langle \boldsymbol{x}, \boldsymbol{b}\right\rangle^2\}.$$

Noting that $\mathrm{var}\{f(\boldsymbol{x})\} = \mathrm{var}\{g(\boldsymbol{x})\} = \mathrm{E}\{g(\boldsymbol{x})^2\} - \mathrm{E}\{g(\boldsymbol{x})\}^2$, the expression for variance in part (b) of the proposition follows from Proposition 2.3 and some algebraic manipulations. □

References

[1] C. S. Withers, "The moments of the multivariate normal," *Bulletin of the Australian Mathematical Society*, vol. 32, no. 1, pp. 103–107, 1985.

[2] B. Holmquist, "Moments and cumulants of the multivariate normal distribution," *Stochastic Analysis and Applications*, vol. 6, no. 3, pp. 273–278, 1988.

[3] K. Triantafyllopoulos, "On the central moments of the multidimensional gaussian distribution," *The Mathematical Scientist*, 2003.