High Dimensional Statistics | Prof. Dr. Podolskij Mark | Homework 1

Anton Zaitsev | 0230981826 | anton.zaitsev.001@student.uni.lu | University of Luxembourg

March 17, 2024

Exercise 1

Let $Y \sim \mathcal{N}_n(\mu, \Sigma), \mu \in \mathbb{R}^n, \Sigma \in \mathbb{R}^{n \times n}$ positive definite, $A \in \mathbb{R}^{n \times n}$.

1. $\mathbb{E}[Y^T AY] = \mu^T A\mu + tr(A\Sigma)$
2. Let A symmetric, then: $cov(Y, Y^T AY) = 2\Sigma A\mu \in \mathbb{R}^n$

Proof (1)

Notice that $\mathbb{E}[Y^T AY]$ is a 1×1 matrix. Then, we can apply the trace operator and get the following:

$$\mathbb{E}[Y^T AY] = tr(\mathbb{E}[Y^T AY])$$

Since trace is a linear operator:

$$tr(\mathbb{E}[Y^T AY]) = \mathbb{E}[tr(Y^T AY)]$$

Since Y and A are $n \times n$ matricies, we can use the cyclic property of the trace operator, i.e. the trace of the product of two square matricies with the same size is invariant under cyclic permutation:

$$\begin{aligned}\mathbb{E}[tr(Y^T AY)] &= \mathbb{E}[tr(YY^T A)] \\ &= \mathbb{E}[tr(AYY^T)]\end{aligned}$$

Now, again, using the fact that the trace operator is linear:

$$\mathbb{E}[tr(AYY^T)] = tr(\mathbb{E}[AYY^T])$$

Since A is deterministic:

$$tr(\mathbb{E}[AYY^T]) = tr(A\mathbb{E}[YY^T])$$

Now, notice that:

$$\Sigma = cov(Y, Y^T) = \mathbb{E}[YY^T] - \mathbb{E}[Y]\mathbb{E}[Y^T] = \mathbb{E}[YY^T] - \mu\mu^T$$

Thus, $\mathbb{E}[YY^T] = \Sigma + \mu\mu^T$ and we get:

$$\begin{aligned}\mathbb{E}[Y^T AY] &= tr(A\mathbb{E}[YY^T]) \\ &= tr(A(\Sigma + \mu\mu^T)) \\ &= tr(A\Sigma + A\mu\mu^T) \\ &= tr(A\Sigma) + tr(A\mu\mu^T) \\ &= tr(A\Sigma) + tr(\mu^T A\mu)\end{aligned}$$

Since $\mu^T A\mu$ is 1×1 matrix, we finally get:

$$\mathbb{E}[Y^T AY] = tr(A\Sigma) + \mu^T A\mu$$

End of proof.

Proof (2)

By the definition of covariance, we have:

$$cov(Y, Y^TAY) = \mathbb{E}[(Y - \mathbb{E}[Y])(Y^TAY - \mathbb{E}[Y^TAY])]$$

From the first part of this exercise we know that $\mathbb{E}[Y^TAY] = tr(A\Sigma) + \mu^TA\mu$ and due to the fact that $\mathbb{E}[Y] = \mu$ we get:

$$\begin{aligned} cov(Y, Y^TAY) &= \mathbb{E}[(Y - \mathbb{E}[Y])(Y^TAY - \mathbb{E}[Y^TAY])] \\ &= \mathbb{E}[(Y - \mu)(Y^TAY - tr(A\Sigma) - \mu^TA\mu)] \end{aligned}$$

Notice that we can rewrite $Y^TAY - \mu^TA\mu$ as follows:

$$\begin{aligned} Y^TAY - \mu^TA\mu &= (Y - \mu + \mu)^TA(Y - \mu + \mu) - \mu^TA\mu \\ &= (Y - \mu)^TA(Y - \mu) + (Y - \mu)^TA\mu \\ &\quad + \mu^TA(Y - \mu) + \mu^TA\mu - \mu^TA\mu \\ &= (Y - \mu)^TA(Y - \mu) + 2(Y - \mu)^TA\mu \end{aligned}$$

Thus, we get:

$$\begin{aligned} cov(Y, Y^TAY) &= \mathbb{E}[(Y - \mathbb{E}[Y])(Y^TAY - \mathbb{E}[Y^TAY])] \\ &= \mathbb{E}[(Y - \mu)(Y^TAY - tr(A\Sigma) - \mu^TA\mu)] \\ &= \mathbb{E}\left[(Y - \mu)\left((Y - \mu)^TA(Y - \mu) + 2(Y - \mu)^TA\mu - tr(A\Sigma)\right)\right] \\ &= \mathbb{E}\left[(Y - \mu)(Y - \mu)^TA(Y - \mu)\right] + 2\mathbb{E}\left[(Y - \mu)(Y - \mu)^T\right]A\mu \\ &\quad - \mathbb{E}\left[(Y - \mu)tr(A\Sigma)\right] \end{aligned}$$

Let us study each of these terms separately. Notice that the first term relates to the third moment of **centered** multivariate normal random variable. By Isserlis' theorem, which states that the **odd** moment of a centered normal random variable is 0, we can deduce that this term is equal to 0. This result also comes from the fact that $Y - \mu \overset{\mathcal{L}}{=} \mu - Y$ (have the same distribution), which implies that $\mathbb{E}[Y - \mu] = \mathbb{E}[-(Y - \mu)] = -\mathbb{E}[Y - \mu] = 0$.
The expectation $2\mathbb{E}\left[(Y - \mu)(Y - \mu)^T\right]$ (second term without $A\mu$) can be expanded as follows:

$$\begin{aligned} 2\mathbb{E}\left[(Y - \mu)(Y - \mu)^T\right] &= 2\left(\mathbb{E}[YY^T] - \mathbb{E}[Y]\mu^T - \mu\mathbb{E}[Y^T] + \mu\mu^T\right) \\ &= 2\left(\mathbb{E}[YY^T] - \mu\mu^T - \mu\mu^T + \mu\mu^T\right) \end{aligned}$$

Notice that

$$\begin{aligned} \mathbb{E}[YY^T] &= cov(Y) + \mathbb{E}[Y]\mathbb{E}[Y^T] \\ &= \Sigma + \mu\mu^T \end{aligned}$$

Thus:

$$\begin{aligned} 2\mathbb{E}\left[(Y - \mu)(Y - \mu)^T\right] &= 2\left(\mathbb{E}[YY^T] - \mathbb{E}[Y]\mu^T - \mu\mathbb{E}[Y^T] + \mu\mu^T\right) \\ &= 2\left(\Sigma + \mu\mu^T - \mu\mu^T - \mu\mu^T + \mu\mu^T\right) \\ &= 2\Sigma \end{aligned}$$

For the third term we have:

$$\mathbb{E}\left[(Y - \mu)tr(A\Sigma)\right] = \mathbb{E}\left[(Y - \mu)\right]tr(A\Sigma)$$

Again, notice that $\mathbb{E}\left[(Y-\mu)\right]$ relates to the first moment of centered multivariate normal variable, which is again 0. Thus, we finally get the following equality:

$$\begin{aligned}
cov(Y, Y^T AY) &= \mathbb{E}[(Y - \mathbb{E}[Y])(Y^T AY - \mathbb{E}[Y^T AY])] \\
&= \mathbb{E}[(Y-\mu)(Y^T AY - tr(A\Sigma) - \mu^T A\mu)] \\
&= \mathbb{E}\left[(Y-\mu)\left((Y-\mu)^T A(A-\mu) + 2(Y-\mu)^T A\mu - tr(A\Sigma)\right)\right] \\
&= \mathbb{E}\left[(Y-\mu)(Y-\mu)^T A(A-\mu)\right] + 2\mathbb{E}\left[(Y-\mu)(Y-\mu)^T A\mu\right] \\
&\quad - \mathbb{E}\left[(Y-\mu)tr(A\Sigma)\right] \\
&= 0 + 2\Sigma A\mu + 0 \\
&= 2\Sigma A\mu
\end{aligned}$$

End of proof.