Exercise 3

Machine Learning I

|  |  |
| --- | --- |
|  | 2A-1. |

Short way:

|  |
| --- |
| Prerequisites |

|  |
| --- |
| The simplest most high-level way is to use predefined rules for matrix differentiation. The rules for this algebra are laid out in the “Matrix Cookbook” on page 8 and page 10.  The book can be found here:  <https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf>  Some important rules form this book (let be a column vector): |

We have:

As usual, we take the logarithmic transform:

Calculation of the mean:

First, let us calculate out the brackets in the previous exponent:

apply the rules of matrix differentiation algebra:

Now we can solve for

Calculation of Variance:

This one is more involved. But by just utilizing the rules of the Matrix Cookbook, we quickly get to a solution. Note: Now I use instead of for convenience.

Now solve for

It appears that the original task that says

is wrong, as the is also part of other solutions found on the web:

<https://stats.stackexchange.com/questions/351549/maximum-likelihood-estimators-multivariate-gaussian>

I have to check that with Bertschinger.

|  |  |
| --- | --- |
|  | 2A-3. |

Direct calculation:

.

Calculation of the exponent and completing the square:

Let

Now insert into the original equation:

|  |  |
| --- | --- |
|  | 2A-4. |

As usual, we are taking the partial derivate :

If we tried to solve this for , we would encounter dependencies on the other . This is an indicator that it would make sense to convert the equation into matrix form and treat the entire derivative as a solution to a system of equations.

Converting each term step by step leads to:

where

is the design matrix. For the entire vector, this would lead to:

Solving for and utilizing the Penrose inverse :