

Assignment 2: Have you been redpilled?

Due date: **10. 12. 23:59** (moved the deadline to give you more time)

Where to turn in: IS folder “Assignment 2” as PDF

The first assignment was all about fancy words like ontology and latent construct. Now, we finally get into the good stuff as this one is all about linear algebra. No philosophy allowed.

To succeed, you are supposed to have conceptual grasp of the concepts we covered in the lecture. That includes, but is not limited to, the basics of matrix operations, types of matrices, linear dependence, and the eigendecomposition of a matrix. Most tasks are purely conceptual and can be done without any software. Computational tasks, on the other hand, are best done in R but, if you want, you can use [WolframAlpha](#) as well. This lets you circumvent the need for learning R. Switch to “Math Input” to input matrices directly. If you elect to use R, I uploaded the tutorial script I showed you in class which might help you.

Also, sorry for the delay! Of course, I moved the deadline, so you have enough time to do this assignment comfortably.

Mandatory:

Mulaik, Fundamentals of Factor Analysis, **Chapter 2**

YouTube series on matrix algebra:

https://www.youtube.com/watch?v=fNk_zzaMoSs&list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab

Optional:

Free online course on linear algebra:

<https://www.khanacademy.org/math/precalculus/x9e81a4f98389efdf:matrices/>

Linear algebra in R:

<https://www.math.uh.edu/~jmorgan/Math6397/day13/LinearAlgebraR-Handout.pdf>

Linear regression via matrix algebra

<https://www.stat.purdue.edu/~boli/stat512/lectures/topic3.pdf>

* * *

Q1: What conclusions can you draw about the following vectors / matrices (i.e., their relationship or properties)?

I. $\mathbf{a} \cdot \mathbf{b} = 0$

II. $\cos(\angle \mathbf{c}\mathbf{d}) = .707$

III. $\mathbf{e} = 2\mathbf{f}$

IV. $\mathbf{M}\mathbf{g} = 3.441\mathbf{g}$

V. $\mathbf{W}\mathbf{X} = \mathbf{X}\mathbf{W} = \mathbf{I}$

(2.5)

Q2: Fill in the blank elements of the following matrix \mathbf{M} so that it's singular. Explain your strategy. (0.5)

$$\mathbf{M} = \begin{matrix} & & 1 \\ & & -2 \\ & & 7 \end{matrix}$$

Q3: “Here’s der data you asked für, Herr Commandant,” you say, struggling to sound as natural as possible. It feels like you’ve been successful after all, yet a slight squeak could have been heard in your voice if one cared to listen close enough. But that’s okay. You try not to be too hard on yourself. It’s not every day you try to deceive Nazi Germany. You are allowed one minuscule squeal.

Your boss glances over the dataset. “I have another job for you.” Luckily, it appears he didn’t care about the specifics of your vocalization. You fail to withhold a relieved sigh. *Tais toi*, you fool!

The commander lifts his head and witnesses your inner monologue. It only takes a second but it’s a second that feels almost as large as the imperial ambitions of your employer. Maybe even larger. During that incredibly imploded eon you enter a state of hypervigilance that lets you – no, *makes* you – notice every detail around including the commander’s nose-hair popping out of his otherwise well-groomed appearance so confidently that it almost seems like a routine part of a Wehrmacht officer’s formal attire. Meine Frau, do I have everything I need? My uniform, my gun, is my nose-hair popping out visibly enough?

Finally, the second passes, as all seconds do, and the commander, of course, does not notice anything. This time, you contain your relief, and respond:

“Yes, sir.”

The authoritative figure in front of you hands you a flashlight from his three-dimensional desk. “It’s regression time baby.”

Your goal is to predict the daily consumption of supplies in several German cities from their population, average temperature and whether they are being shelled by the Allies. The army has lost all its software licenses, so you need to do it by hand. Don’t forget to interpret the results.

Oh, and you are free to mislead the Germans, just make sure you do it in a smart way, so you don’t get caught! (2)

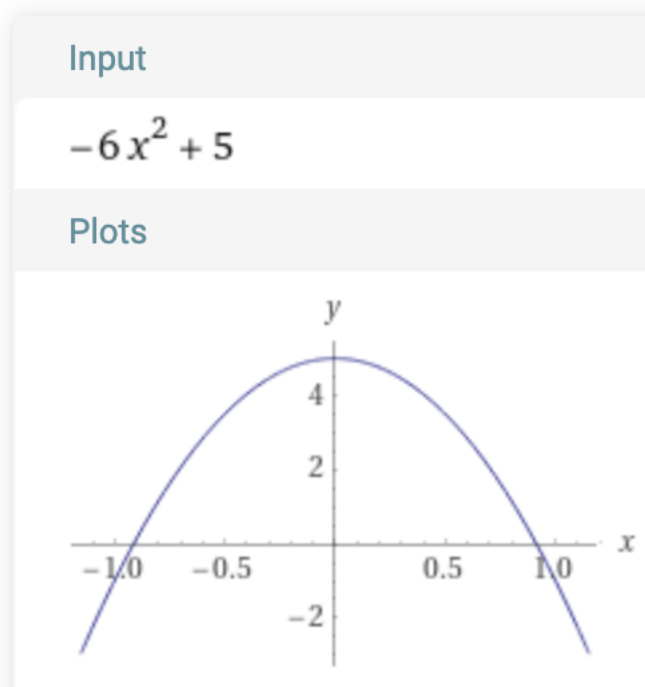
33
28
daily supply consumption = 12
45
51

1
-1
avg. temperatures = 2
0
1

340500
240000
populations = 50000
280000
320000

1
1
shellings = 0
1
0

Q4: Assume the following figure is the likelihood function of a parameter x . The value of the function (y) represents the likelihood of the specific values of the parameter x . What is the maximum likelihood estimate for x ? How did you tell? You can use, e.g., WolframAlpha to help you answer. (1.5)



Q5: The cracking of a bone being broken breaks the silence in the basement you and your partner spent the last six hours being confined in. It's not your first day undercover investigating the Mafia ring – the sounds of others' fractures stopped moving you a while ago. But still, when it is your own limb being broken, you can't help but to feel *something*.

The goons are, however, disappointed in your lack of emotion. For a while, they stare at you blankly but quickly shapeshift back into their grinning personas.

“You may have heard of the Russian roulette, right,” one of them grins even grimmer. “But that's not something we do, don't worry.” He pauses, trying to sound dramatic and make you worry. “That's because we play the *Gaussian* roulette here.”

The goon presents you six correlation matrices. One of them is faked, five of them actually computed from real-world data. Can you save yourself and your partner by picking the fake one

and winning the Gaussian roulette? Assume you were provided population correlation matrices.
(1)

$$R_1 = \begin{matrix} & 1 & .30 & .44 \\ & .30 & 1 & .34 \\ & .44 & .34 & 1 \end{matrix}$$

$$R_2 = \begin{matrix} & 1 & .49 & -.34 \\ & .49 & 1 & .45 \\ & -.34 & .45 & 1 \end{matrix}$$

$$R_3 = \begin{matrix} & 1 & .73 & -.70 \\ & .73 & 1 & .72 \\ & -.70 & .72 & 1 \end{matrix}$$

$$R_4 = \begin{matrix} & 1 & .34 & .21 \\ & .34 & 1 & .33 \\ & .21 & .33 & 1 \end{matrix}$$

$$R_5 = \begin{matrix} & 1 & .015 & .843 \\ & .015 & 1 & .550 \\ & .843 & .550 & 1 \end{matrix}$$

$$R_6 = \begin{matrix} & 1 & .19 & .45 \\ & .19 & 1 & .33 \\ & .45 & .33 & 1 \end{matrix}$$

Q6: If a variance-covariance matrix of scale items is a diagonal matrix, what is the Cronbach's Alpha of this scale? (0.5)

Q7: Explain why a variance-covariance matrix has variances on the diagonal. (0.5)

Q8: How many eigenvalues of square matrix \mathbf{M} can you guess exactly if you know it is singular? Why? (0.5)

Q9: In ordinary least squares (OLS) regression we find the solution by solving the equation

$$\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{Y}$$

where \mathbf{X} is a $n \times p$ matrix representing your predictors (n people and p predictor variables), $\hat{\boldsymbol{\beta}}$ (read "beta hat") is a $p \times 1$ matrix of regression coefficients (one for each predictor) and \mathbf{Y} is a $n \times 1$ matrix of each person's values on the dependent variable.

The solution goes like this:

$$\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{Y}$$

First, we premultiply both sides by the transpose of \mathbf{X}

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

And then we keep the regression weight matrix on one side and the rest on the other by multiplying both sides by the inverse of $\mathbf{X}'\mathbf{X}$.

Why not just pre-multiply with \mathbf{X}^{-1} right away? (1)