

Math homework solutions with detailed steps

Talon Meyer • talon_meyer@berkeley.edu



File: hw0.pdf

I attached the HW0 pdf with no modifications and gave Claude a simple prompt asking for solutions to the mathematics questions. Spoiler alert: Claude was able to solve these questions in one shot with no errors, which frankly surprised me.

Solve the math sections (questions 2, 3, 4, 5) of the attached homework assignment. Show your work in detail and carefully explain each step you make.

12/8/2025, 8:47:18 PM

Thought process: Mapped systematic approach to solving neural network mathematics problems. I find it interesting that Claude labels it's thought process as thought process.

The user wants me to solve the math sections (questions 2, 3, 4, 5) of this homework assignment on Deep Neural Networks. Let me work through each problem carefully.

Describes what it will do, aligns with classic CoT reasoning theory.

Let me start by understanding what's being asked in each section:

Question 2: Vector Calculus Review

- Derivatives of vector expressions
- Convention: vector derivatives of scalar w.r.t. column vector are row vectors

Question 3: Least Squares and Min-norm from SVD perspective

- Overdetermined and underdetermined systems
- Using SVD to derive solutions

Question 4: 5 Interpretations of Ridge Regression

- Various perspectives on ridge regression

Question 5: ReLU Elbow Update under SGD

- Understanding ReLU behavior with gradient descent

Let me solve these step by step and create a comprehensive document. Claude Opus attempts to read and understand what each question is asking before it begins solving. I find this fascinating considering I also do this and I know many of my peers follow this practice as well.

I'll solve all the math sections of this homework assignment in detail. Let me create a comprehensive document with all the solutions. [This section is actually displayed to the user.](#)

Create File: Creating a detailed solution document for the math sections (Questions 2-5) of the EECS 182 Homework 0 [Describes the agentic actions its taking and tools its using.](#)

Thought process: Assembled detailed solutions and organized deliverables.

I've created a comprehensive markdown file with all the solutions. Now let me copy it to the outputs folder so the user can access it. [See the markdown file under the chat history PDF.](#)

Bash Tool: Copy the solutions file to outputs for user access

[Again describing its tool usage, I find it remarkable how proficient Claude has gotten at using programming tools like Bash.](#)

Thought process: Organized mathematical solutions and prepared detailed explanations.

The file is now in the outputs directory. Let me provide a summary of the solutions I created.

I've solved all the math sections (Questions 2-5) of the homework with detailed explanations. Here's a summary of what's covered:

[Claude reports to the user it finished its given task. It's worth mentioning that before this I was able to watch as Claude wrote the markdown file containing the HW solutions.](#)

Question 2: Vector Calculus Review

In the following sections Claude summarizes to the user its findings and how it solved each question. The details are shown in the generated markdown file.

I find it interesting that the actual chat uses different characters than in its solutions. In the solution Claude used LaTex code embedded within a md file, but here it uses special characters when possible and does its best when not.

- **(a)** Showed $\partial/\partial x(x^T c) = c^T$ by expanding the inner product and differentiating term by term
- **(b)** Showed $\partial/\partial x||x||^2_2 = 2x^T$ using the squared norm expansion
- **(c)** Showed $\partial/\partial x(Ax) = A$ by computing the Jacobian matrix
- **(d)** Showed $\partial/\partial x(x^T Ax) = x^T(A + A^T)$ by carefully handling the quadratic form
- **(e)** The derivative equals $2x^T A$ when **A is symmetric** ($A = A^T$)

I find it curious that in the above section is detailed each subpart with (a), (b), ... while in the next sections it does not. I suppose this speaks to the inherent non-determinism within these models.

Question 3: Least Squares & Min-norm via SVD

- Derived the **least squares solution** $w^* = (X^T X)^{-1} X^T y$ for overdetermined systems
- Derived the **min-norm solution** $w^* = X^T(XX^T)^{-1} y$ for underdetermined systems
- Showed both simplify to $w^* = V\Sigma U^T y$ using SVD
- Explained left-inverse and right-inverse properties

I find it interesting that Claude described its process in this section as opposed to the answers. I would think the answers would be the more relevant information. Perhaps Claude is trained to summarize what it did as opposed to provide concise answers when extended thinking is enabled.

Question 4: Five Perspectives on Ridge Regression

- **Optimization:** Derived $w = (X^T X + \lambda I)^{-1} X^T y$ using calculus
- **SVD perspective:** Showed ridge "shrinks" small singular values while preserving large ones
- **MAP estimation:** Proved ridge is the MAP estimate with Gaussian prior and noise
- **Fake data:** Showed augmenting data with $\sqrt{\lambda}I_d$ yields identical solution
- **Fake features:** Showed min-norm on augmented features gives ridge solution
- Proved equivalence of the two ridge formulas using matrix identities
- Analyzed $\lambda \rightarrow \infty$ (shrinkage to zero) and $\lambda \rightarrow 0$ (OLS/min-norm limits)

Question 5: ReLU Elbow Updates

- Found elbow location $e = -b/w$
- Derived all relevant gradients for single ReLU
- Analyzed how gradient descent affects slope and elbow position in each case
- Extended to full network: $e_i' = -(b_i - \lambda\delta)/(W_i^{(1)} - \lambda\delta x)$ where δ captures the backpropagated error

[View your solutions document](#)

12/8/2025, 8:50:04 PM

Again here Claude uses odd mathematical notation. It's as if it expects this to be compiled into LaTex. This could be a touch of laziness on the side of the training engineers who may have trained Claude to output solutions in LaTex whether or not it will compile. However, it also uses the ASCII representations of Greek letters whenever possible, so perhaps not.

