CS5242 : Neural Networks and Deep Learning

Administrative (Week 5)

Semester 1 2021/22

Xavier Bresson

https://twitter.com/xbresson

Department of Computer Science
National University of Singapore (NUS)

Tentative Schedule

Week	Topic	Assessment
1	Introduction	
2	Vanilla NNs - Part 1	
3	Vanilla NNs - Part 2	
4	MLP - Part 1	
5	MLP - Part 2	
6	CNNs - Part 1	Quiz 1
Recess		
7	CNNs - Part 2	Coding test 1
8	RNNs - Part 1	
9	RNNs - Part 2	Quiz 2
10	ANNs - Part 1	
11	ANNs - Part 2	Coding test 2
12	Conclusion	
13		Project delivery

NNs = Neural Networks
MLP = Multi-Layer Perceptron
CNNs = Convolutional Neural Networks
RNNs = Recurrent Neural Networks
ANNs = Attention Neural Networks

Forum

Questions on Lecture 5 Slide 21

Posted by **LAW ANN LIAT LARRY** on 5 Sep 2021 11:23 am. Last modified on 5 Sep 2021 11:26 pm.

Hi, I've a few questions on Lecture 5 Slide 21.

1. The example has a batch size of 2, which implies that the gradient tensor dL/dW needs to be averaged (lecture 4, slide 40), yet the updating of weights dL / dW = EZ^T has no indication of such averaging. In fact, dL / dW = EZ^T = [EZ^T_1 | EZ^T_2], which implies that the ith signal updates the ith column of the weights. Where did my understanding go wrong?

2. As a follow-up question, does the input signal X in slide 19 include batch size?

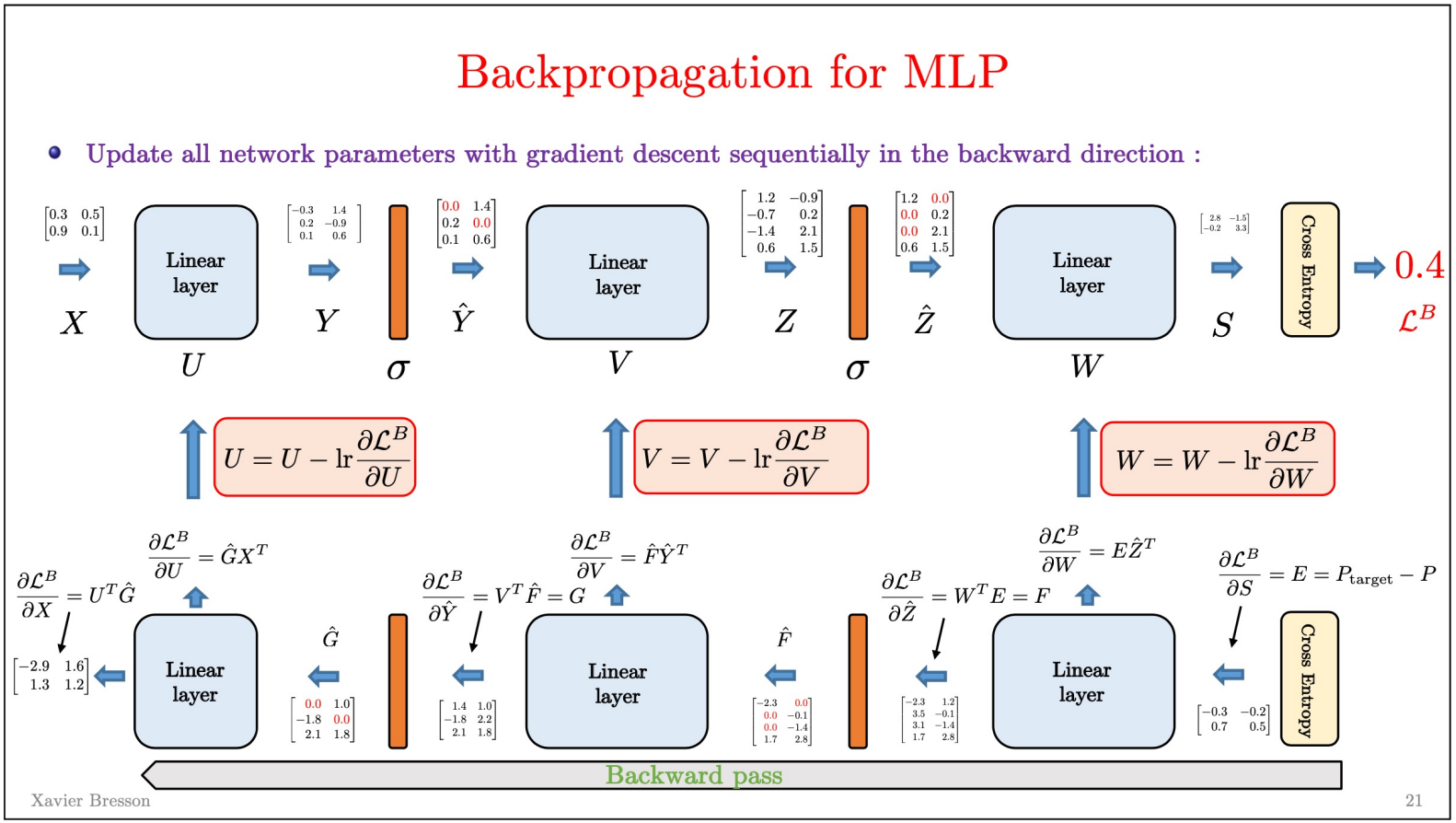
3. Applying formula of slide 22 to the last ReLU layer of slide 21, dL / dZ = step(Z) x F is an illegal operation as step(Z) is of shape (4, 2) and F is of shape (4, 2). Yong Liang explained to me that step(Z) x F is point wise multiplication. If so, how does the previous step that involves matrix multiplication (i.e. (dY / dX)^T • dL / dY) lead to the next step that now involves point-wise multiplication (i.e. step(Z) x F)? Where did my understanding go wrong?

If possible, I'd prefer concrete calculus examples as I have trouble understanding the abstract notations.

Thank you!

Forum

- Slide 21 Lecture 5:



Forum

Question 1 :

$$\underbrace{\frac{\partial L}{\partial W}}_{2\times 4} = \underbrace{E}_{2\times 2}\underbrace{\vec{Z}^T}_{(4\times 2)^T \to 2\times 4}$$

$\underbrace{\quad}_{2\times 4}$ $\qquad$ $\underbrace{\quad}_{2\times 4}$

General case :

$$\underbrace{\frac{\partial L}{\partial W}}_{2\times 4} = \underbrace{E}_{2\times n}\underbrace{\vec{Z}^T}_{n\times 4}$$

n ↖↗ batch size

+ usually $L = \frac{1}{n}$ <cross-entropy>

= mean operation

Question 2 : Yes

Forum

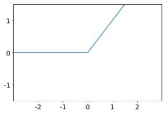
- Slide 22 Lecture 5:

Backpropagation for ReLU layer

- **Forward pass:**

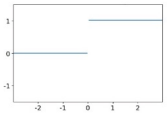
Input signal X → **ReLU layer** → $Y = \max(X, 0)$ Output signal

No weight parameter !



- **Backward pass:**

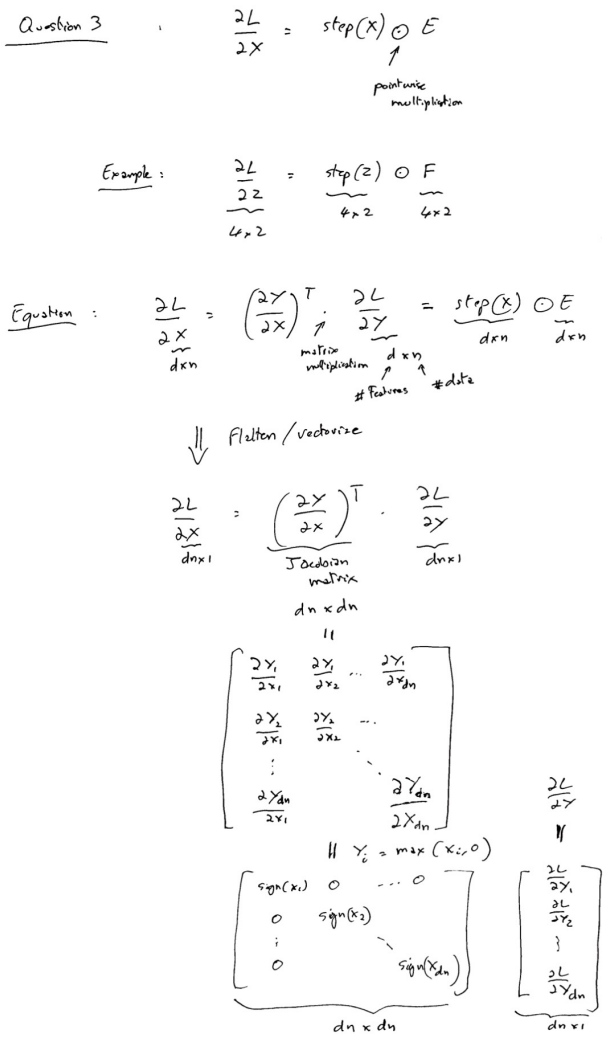
$$\frac{\partial Y}{\partial X} = \frac{\partial}{\partial X}\max(X,0) = \begin{cases} 1 & \text{if X value is positive} \\ 0 & \text{otherwise} \end{cases} = \text{step}(X)$$



Output gradient $\frac{\partial \mathcal{L}^B}{\partial X} = \left(\frac{\partial Y}{\partial X}\right)^T \cdot \frac{\partial \mathcal{L}^B}{\partial Y}$ $= \text{step}(X).E$ ← **ReLU layer** ← $E = \frac{\partial \mathcal{L}^B}{\partial Y}$ Input gradient

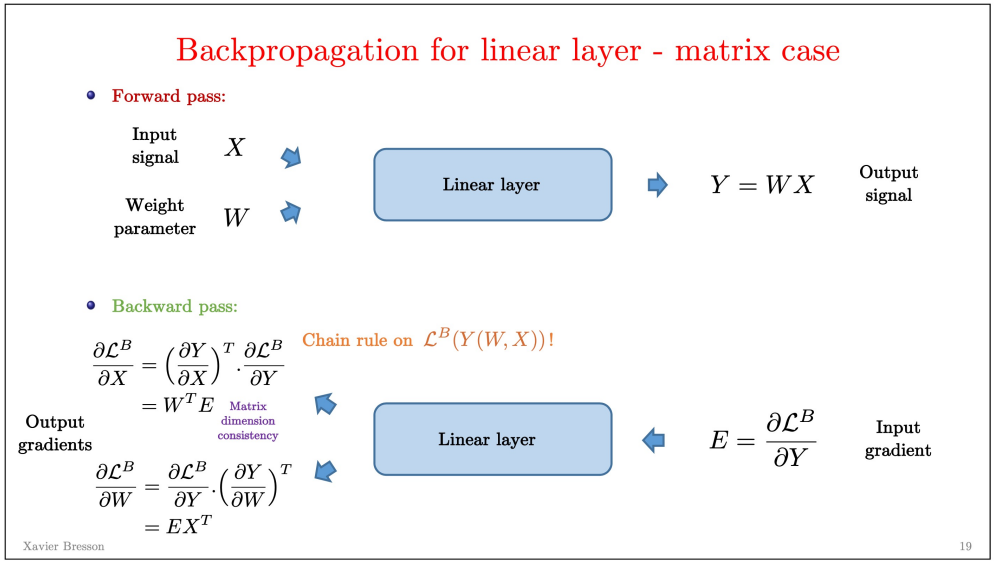
Chain rule on $\mathcal{L}^B(Y(W,X))$!

Forum


Question 3
$\frac{\partial L}{\partial X} = step(X) \odot E$
pointwise multiplication
Example:
$\frac{\partial L}{\partial Z} = step(Z) \odot F$
4×2 4×2 4×2
Equation:
$\frac{\partial L}{\partial X} = \left(\frac{\partial Y}{\partial X}\right)^T \cdot \frac{\partial L}{\partial Y} = step(X) \odot E$
$d \times n$ $d \times n$ $d \times n$
matrix multiplication
$d \times n$
Features
data
Flatten / vectorize
$\frac{\partial L}{\partial X} = \underbrace{\left(\frac{\partial Y}{\partial X}\right)^T}_{\text{Jacobian matrix}} \cdot \frac{\partial L}{\partial Y}$
$dn \times 1$ $dn \times dn$ $dn \times 1$
$dn \times dn$
$\frac{\partial L}{\partial Y}$
$Y_i = max(X_i, 0)$
$sign(x_1)$ $sign(x_2)$ $sign(x_{dn})$
$dn \times dn$
$dn \times 1$

Forum

- Slide 19 Lecture 5:
  - http://cs231n.stanford.edu/handouts/linear-backprop.pdf
  - More generally about matrix calculus
    - https://en.wikipedia.org/wiki/Matrix_calculus
    - The Matrix Calculus You Need For Deep Learning, https://arxiv.org/pdf/1802.01528.pdf



CS5242 Evaluation

- This module is 100% CA, i.e., there is no final exam.
- There are 3 components:
  - 2 quizzes, each 15% (Weeks 6 & 9). The quizzes are individual. There is no makeup quiz. The weight of quizzes is 30% in total.
  - 2 coding tests, each 20% (Weeks 7 & 11). The coding tests are individual. Check the schedule for the date of each coding test. There is no makeup coding test. The weight of quizzes is 40% in total.
  - 1 group project, 30% (Week 13). Check the schedule for the date of the project delivery. The project is group-based with a group size of at most 3. Choose your group wisely -- each teammate must contribute equally to the project. Each project will deliver a python notebook with the code and the description of the project (in Markdown), and a short video presentation (each student will present her/his contribution to the project).
    - I will introduce the project on Week 6.

Quiz (Week 6)

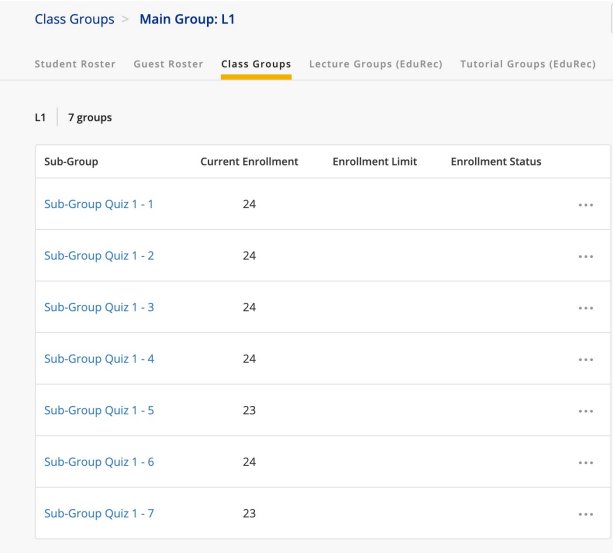
- Instructions :
  - Time: 1hr/60min on Tue Sept 14^{th} 8:10pm-9:10
  - The test will cover all material up to Lecture 6 "MLP - Part 2" included.
  - The test has 10 questions, from easy, intermediate to difficult questions.
    - No coding questions
  - Open-book exam : You can use any material (internet, etc).
  - Questions will focus on the understanding of the fundamental concepts of deep learning techniques.

Quiz (Week 6)

- Instructions :
  - The 10 questions will be provided in LumiNUS => Quiz => Quiz1
    - The quiz will be open at 8:10pm.The password will be available at 8:08pm via LumiNUS => Announcements.
    - You will be able to submit your answers up to 9:10pm, then the system will automatically close the quiz at 9:10pm.
  - Individual test
    - We will use online invigilation.
    - Two devices will be needed :
      - Phone/tablet : Use to record your face and your background with Mobile Zoom.
      - Laptop/desktop computer : Use to record your screen with Desktop Zoom.
      - Slides for setup instructions are available at https://drive.google.com/drive/u/0/folders/1LJeeG4CQ4Mea-Nvt5fAXPkKYKxaWEINx, file : "setup_online_invigilation.pdf"
      - Important : Do not make the setup of the devices on the day of the test !
        - If any issue with the setup, contact the TA in charge of your group (see next slide) before the day of the test.

Quiz (Week 6)

- Instructions :
  - You are assigned to one group, LumiNUS=>Class & Groups=>ClassGroups=>L1=>Groups Quiz 1.
    - Note that groups are randomly allocated and you will have a new group at each test.
  - Each group will be invigilated (zoom link below) and graded by an assigned TA (see next slide).



Quiz (Week 6)

- Instructions :
  - TA assigned to the groups :
    - Group 1 : Mr Wu Zhaomin, zhaomin@u.nus.edu, https://nus-sg.zoom.us/j/6346520585?pwd=VlRUMVArKzE2UjZNVGRyRkpCS1JWdz09
    - Group 2 : Mr Hu Sixu, husixu@u.nus.edu, https://nus-sg.zoom.us/j/2256639902?pwd=TDZMTlMrbTVFd3ZUU2NIc1RtMFBldz09
    - Group 3 : Mr Goh Yong Liang, gyl@u.nus.edu, https://grab.zoom.us/j/9970260265?pwd=NE1xeGVNekNsZFk1VVhlV0h3cUFxZz09
    - Group 4 : Mr Wang Guangzhi, guangzhi.wang@u.nus.edu, https://nus-sg.zoom.us/j/3731446480?pwd=ZGpuWVQ4YmtCbDBhcC9hQmVic3R5UT09
    - Group 5 : Mr Fu Yujian, e0427770@u.nus.edu, https://nus-sg.zoom.us/j/9880635334?pwd=QnZkdkRVZ1NIMFk5a1VhRFBEU1AvQT09
    - Group 6 : Mr Liu Hongfu, e0673183@u.nus.edu, https://nus-sg.zoom.us/j/6502616852?pwd=ZXA5SWtITHp4dHF5T1NMdDF6OGxWdz09
    - Group 7 : Mr Liu Xu, liuxu12@u.nus.edu, https://nus-sg.zoom.com.cn/j/3227470602?pwd=YWVsRmY3bTQvdjMvVUJKSUg0NGpUZz09

Quiz (Week 6)

- Schedule of the quiz :
  - You must login/join your zoom group (zoom link provided in the previous slide) at 7:50pm.
  - You must start the zoom recording of your face and your background with your phone/tablet between 8:00pm-8:10, see slides "setup_online_invigilation.pdf".
  - Your TA invigilator will make the attendance between 7:50pm-8:10.
    - Present your student ID to the TA when your name is asked.
  - You will receive the password of the quiz at 8:08pm in LumiNUS=>Announcements.
  - The Quiz will be open at LumiNUS=>Quiz=>Quiz1 and will last 60min from 8:10pm-9:10. Then the system will automatically close the quiz at 9:10pm.
  - Note: Only the first attempt will be marked. DO NOT join the quiz multiple times.
  - You have until midnight of the same day to upload your video of you and your background to LumiNUS=>Files=>Quiz Screen Video Upload=>Quiz 1, see slides "setup_online_invigilation.pdf".

- Class at 6:30pm :
  - I will present the project (30min).

Quiz (Week 6)

- Grade release :
  - We will release the grade of the quiz on Tue Sept 21st in LumiNUS.
  - If you have any question regarding the quiz, contact the TA in charge of your group.
    - Please, be polite and respectful when discussing the quiz.

Questions?