ממן 13 – שאלה 2

שלומי דומננקו

נתון:

- תמונה X: 1x3x3
- גרעין K: 1x2x1
- $Y = conv(X, K)$
- ידוע לנו: $\frac{\delta C}{\delta y_{i,j}}$ לכל i,j

מטרה:

למצוא ביטוי ל $\frac{\delta C}{\delta K}$ וגם $\frac{\delta C}{\delta X}$

נסמן:

$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,1} & x_{3,2} & x_{3,3} \end{pmatrix}$$

$$K = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$$

אנו יודעים כי הקרנל K יכול להיכנס ל X שלוש פעמים בעמודות ופעמיים בשורות. ולכן:

$$Y = \begin{pmatrix} y_{1,1} & y_{1,2} & y_{1,3} \\ y_{2,1} & y_{2,2} & y_{2,3} \end{pmatrix}$$

נסמן:

בשורה הראשונה:

$$y_{1,1} = x_{1,1}k_1 + x_{2,1}k_2$$

$$y_{1,2} = x_{1,2}k_1 + x_{2,2}k_2$$

$$y_{1,3} = x_{1,3}k_1 + x_{2,3}k_2$$

בשורה השנייה:

$$y_{2,1} = x_{2,1}k_1 + x_{3,1}k_2$$

$$y_{2,2} = x_{2,2}k_1 + x_{3,2}k_2$$

$$y_{2,3} = x_{2,3}k_1 + x_{3,3}k_2$$

כלומר יש לנו ביטוי לכל Y.

ביטוי כללי:

$$y_{i,j} = x_{i,j}k_1 + x_{i+1,j}k_2$$

נמצא ביטוי כעת למה שנתבקשנו.

$$\frac{\delta C}{\delta K}=\sum_{n=1}^{2}\sum_{m=1}^{3}\frac{\delta C}{\delta y_{n,m}}*\frac{\delta y}{\delta k_{1,2}}=\sum_{n=1}^{2}\sum_{m=1}^{3}\frac{\delta C}{\delta y_{n,m}}*\frac{\delta(x_{n,m}k_1+x_{n+1,m}k_2)}{\delta k_{1,2}}=\boxed{\sum_{n=1}^{2}\sum_{m=1}^{3}\frac{\delta C}{\delta y_{n,m}}*\begin{pmatrix}x_{n,m}\\x_{n+1,m}\end{pmatrix}}$$

כלומר זה ביטוי שהוא ווקטור מגודל 2x1.

נזכור שנגזרת של $x_{n,m}k_1$ לפי k_1 הוא כמובן $x_{n,m}$ (כי זה נחשב כקבוע).

נזכור שנתון לנו: $\frac{\delta C}{\delta y_{i,j}}$ (אפשר לראות בעמוד הראשון), ולכן אי אפשר לצמצם יותר.

כעת, נמצא ביטוי לנגזרת של ה LOSS FUNCTION לפי X (הפיקסלים).

$$\frac{\delta C}{\delta X}=\begin{pmatrix}\frac{\delta C}{\delta x_{1,1}}\\\frac{\delta C}{\delta x_{1,2}}\\\frac{\delta C}{\delta x_{1,3}}\\\frac{\delta C}{\delta x_{2,1}}\\\vdots\\\frac{\delta C}{\delta x_{3,3}}\end{pmatrix}$$

נמצא ביטוי לכל אחד מן האיברים של ווקטור זה.

1)
$$\frac{\delta C}{\delta x_{1,1}}=\frac{\delta C}{\delta y_{1,1}}*\frac{\delta y}{\delta x_{1,1}}=\frac{\delta C}{\delta y_{1,1}}*k_1$$

למה? כי:

$$\frac{\delta y_{1,1}}{\delta x_{1,1}}=(x_{1,1}k_1+x_{2,1}k_2)*\frac{1}{\delta x_{1,1}}=k_1$$

קבוע

אפשר לראות שזה נגזרת פשוטה. נמשיך לחשב:

2)
$$\frac{\delta C}{\delta x_{1,2}}=\frac{\delta C}{\delta y_{1,2}}*\frac{\delta y}{\delta x_{1,2}}=\frac{\delta C}{\delta y_{1,2}}*\frac{x_{1,2}k_1+x_{2,2}k_2}{\delta x_{1,2}}=\frac{\delta C}{\delta y_{1,2}}*k_1$$

3)
$$\frac{\delta C}{\delta x_{1,3}}=\frac{\delta C}{\delta y_{1,3}}*\frac{\delta y}{\delta x_{1,3}}=\frac{\delta C}{\delta y_{1,3}}*\frac{x_{1,3}k_1+x_{2,3}k_2}{\delta x_{1,3}}=\frac{\delta C}{\delta y_{1,3}}*k_1$$

נעבור לשורה שנייה:

4) $$\frac{\delta C}{\delta x_{2,1}} = \frac{\delta C}{\delta y_{2,1}} * \frac{\delta y}{\delta x_{2,1}} = \frac{\delta C}{\delta y_{2,1}} * \left(\frac{\delta y_{1,1} + \delta y_{2,1}}{\delta x_{2,1}}\right) = \frac{\delta C}{\delta y_{2,1}} * \frac{(x_{1,1}k_1 + x_{2,1}k_2) + (x_{2,1}k_1 + x_{3,1}k_2)}{\delta x_{2,1}}$$
$$= \frac{\delta C}{\delta y_{2,1}} * (k_2 + k_1)$$

הסיבה שעושים חיבור בין $y_{1,1}, y_{2,1}$ זה בגלל ששניהם משפיעים על אותו איבר: $x_{2,1}$. נצייר ציור שממחיש זאת:

$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,1} & x_{3,2} & x_{3,3} \end{pmatrix}$$

ניתן לראות בבירור שכל שאר ה Yים מתאפסים, ורק $y_{1,1}, y_{2,1}$ משפיעים על הנגזרת של C לפי $x_{2,1}$.

כמו כן ניתר לקצר תהליכים, כי זה מתקיים גם ל $x_{2,2}, x_{2,3}$ כך:

5) $$\frac{\delta C}{\delta x_{2,2}} = \frac{\delta C}{\delta y_{2,2}} * (k_2 + k_1)$$

6) $$\frac{\delta C}{\delta x_{2,3}} = \frac{\delta C}{\delta y_{2,3}} * (k_2 + k_1)$$

כעת נעבור לשורה השלישית:

7) $$\frac{\delta C}{\delta x_{3,1}} = \frac{\delta C}{\delta y_{2,1}} * \frac{\delta y}{\delta x_{3,1}} = \frac{\delta C}{\delta y_{2,1}} * \frac{x_{2,1}k_1 + x_{3,1}k_2}{\delta x_{3,1}} = k_2$$

כבר אפשר לראות שזה יצא אותו דבר גם עבור שאר איברי השורה השלישית. הוכחה:

8) $$\frac{\delta C}{\delta x_{3,2}} = \frac{\delta C}{\delta y_{2,2}} * \frac{\delta y}{\delta x_{3,2}} = \frac{\delta C}{\delta y_{2,2}} * \frac{x_{2,2}k_1 + x_{3,2}k_2}{\delta x_{3,2}} = k_2$$

9) $$\frac{\delta C}{\delta x_{3,3}} = \frac{\delta C}{\delta y_{2,3}} * \frac{\delta y}{\delta x_{3,2}} = \frac{\delta C}{\delta y_{2,3}} * \frac{x_{2,3}k_1 + x_{3,3}k_2}{\delta x_{3,3}} = k_2$$

ולמעשה חישבנו הכל, את כל 9 המשוואות עבור X ואת כל 2 המשוואות עבור K.