Machine Learning for Engineering and Science Applications

Matrix Calculus
(Slightly Advanced)

Motivation

- Machine Learning training requires one to evaluate how one vector changes with respect to another
  - For example how output changes with respect to parameters
- This requires “matrix” calculus
- We will see some initial relations in this video
  - It is useful to understand these, but most of the course can be understood without this portion too.
- More advanced relations exist
  - Suggested resource : https://explained.ai/matrix-calculus/index.html

Scalars and vectors

- $$\left(\frac{\partial \boldsymbol{a}}{\partial x}\right)_i = \frac{\partial a_i}{\partial x}$$

$$\left(\frac{\partial x}{\partial \boldsymbol{a}}\right)_i = \frac{\partial x}{\partial a_i}$$

Vectors and vectors

- $$\left(\frac{\partial \boldsymbol{a}}{\partial \boldsymbol{b}}\right)_{ij} = \frac{\partial a_i}{\partial b_j}$$

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{x} \cdot \boldsymbol{a}) = \frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^T \boldsymbol{a}) = \frac{\partial}{\partial \mathbf{x}}(\boldsymbol{a}^T \mathbf{x}) = \boldsymbol{a}$$

Matrices and vectors

- $$\frac{\partial}{\partial \mathbf{x}}(\boldsymbol{AB}) = \frac{\partial \boldsymbol{A}}{\partial \mathbf{x}}\boldsymbol{B} + \boldsymbol{A}\frac{\partial \boldsymbol{B}}{\partial \mathbf{x}}$$

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^T A\mathbf{x}) = (A + A^T)\mathbf{x}$$

Proof slightly involved (we will see a quick verification on the next slide)

NOTE: For symmetric matrices, $\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^T A\mathbf{x}) = 2A\mathbf{x}$

Derivative of the quadratic form

- $$\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^T A\mathbf{x}) = (A + A^T)\mathbf{x}$$

Verification for 2×2 case