Default Matrix Orientation

Ng Yen Kaow

Inner product notations

- A vector is by default a column
  - For vectors x and y, inner (or dot) product $\langle x, y \rangle = x^{\mathrm{T}} y$
    - $\langle x+z, y \rangle = \langle x, y \rangle + \langle z, y \rangle = x^T y + z^T y$
  - Some texts use row vectors and $\langle x, y \rangle = x y^{\mathrm{T}}$
- For a matrix
  - Each row represents an example or datapoint
  - Each column is a feature
    - Why so? Imagine a spreadsheet
  - For matrices X and Y, $\langle X, Y \rangle = X Y^{\mathrm{T}}$ or $\left(x_i y_i^{\mathrm{T}} \right)$
  - Some texts define columns as examples and use $X^{\mathrm{T}} Y$ as the inner product
- Just write $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}$, $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{x}$, and $\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}$, $\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}$, K?

Inner product notations

- Examples of inner products
  - Given an $n \times m$ matrix X where the rows are datapoints and columns are features
    - The $n \times n$ matrix XX^{T} is the Gram matrix
      - Used in **MDS**
    - The $m \times m$ matrix $X^{\mathrm{T}}X$ is the covariance matrix
      - Used in **PCA**

Outer product notations

- The outer product of two vectors is a matrix

$$\begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} c & d \end{pmatrix} = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix}$$

- The outer product (or Kronecker product) of two matrices is a **tensor**
- We don’t deal with outer products yet, cowboy

Python inner product

- Inner products are performed with `np.dot()`
  - When called on two arrays, the arrays are **automatically** oriented to perform inner product
    - But note that `[[1],[1]]` is a 1×2 matrix
  - When called on an array x and a matrix X, the array is **automatically** read as a row for `np.dot(x,X)`, and column for `np.dot(X,x)` to perform inner product
  - When called on two matrices, make sure that the matrices are oriented correctly, or you will get $X^{\mathrm{T}}X$ when you want XX^{T}
- If you write $x * y$ or $X * Y$, what you get is an element-wise multiplication