Section 1: Matrix Derivatives

CS 205A: Mathematical Methods for Robotics, Vision, and Graphics (Winter 2018)
Stanford University

Matrix Derivatives

This section provides some definitions, derivations, and identities for vector and matrix gradients, which is helpful to understand the image alignment example (4.1.4) from the textbook.

Vector Derivatives

Let f be a scalar function of $\vec{v}$. The gradient of f with respect to $\vec{v}$ is defined:

$$\nabla_{\vec{v}} f(\vec{v}) = \begin{bmatrix} \frac{\partial f}{\partial v_1} \\ \frac{\partial f}{\partial v_2} \\ \vdots \\ \frac{\partial f}{\partial v_n} \end{bmatrix} \tag{1}$$

Using this definition, let's derive some gradients! First, let's try to compute the following gradient (where $\vec{a}$, $\vec{b}$ $\in \mathrm{R}^n$) :

$$\nabla_{\vec{v}} \vec{a}^T \vec{v}$$

If we find an equation for $\frac{\partial f}{\partial v_i}$ (a single component of v), we can fill out the entire gradient using these terms.

$$\frac{\partial \vec{a}^T \vec{v}}{\partial v_i} = \frac{\partial}{\partial v_i}(a_1 v_1 + a_2 v_2 + ... + a_n v_n)$$

$$\frac{\partial \vec{a}^T \vec{v}}{\partial v_i} = a_i$$

Vectorizing this result to the entire gradient:

$$\nabla_{\vec{v}} \vec{a}^T \vec{v} = \vec{a}$$

Next, let's derive:

$$\nabla_{\vec{v}} \vec{v}^T \vec{v}$$

Using a similar logic as before:

$$\frac{\partial \vec{v}^T \vec{v}}{\partial v_i} = \frac{\partial}{\partial v_i}(v_1^2 + v_2^2 + ... + v_n^2)$$

$$\frac{\partial \vec{v}^T \vec{v}}{\partial v_i} = 2 v_i$$

Vectoring this result to the entire gradient:

$$\nabla_{\vec{v}} \vec{v}^T \vec{v} = 2 \vec{v}$$

Matrix Derivatives

Let f be a scalar function of A. The gradient of f with respect to A is defined:

$$\nabla_A f(A) = \begin{bmatrix} \frac{\partial f}{\partial a_{11}} & \frac{\partial f}{\partial a_{12}} & \cdots & \frac{\partial f}{\partial a_{1n}} \\ \frac{\partial f}{\partial a_{21}} & \frac{\partial f}{\partial a_{22}} & \cdots & \frac{\partial f}{\partial a_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial a_{n1}} & \cdots & \cdots & \frac{\partial f}{\partial a_{nn}} \end{bmatrix} \tag{2}$$

Proving identities for matrix derivatives is a considerable amount of algebra, so here I will list some proven identities:

$$\nabla_A \overrightarrow{x}^T A^T A \overrightarrow{x} = 2A \overrightarrow{x} \overrightarrow{x}^T$$

$$\nabla_A \overrightarrow{x}^T A \overrightarrow{y} = \overrightarrow{x} \overrightarrow{y}^T$$

For more details, see section 0.5 of https://cs.nyu.edu/ roweis/notes/matrixid.pdf.