CS7ET01: Mathematics

Dr. Rozenn Dahyot

Tutorial: 18 October 2007

1. **Derivatives.** This tutorial has the purpose to make you feel confortable computing the derivative of an expression defined with vectors and matrices, w.r.t. a vector.

   (a) Show that
   $$\frac{\partial(\mathrm{A}\mathbf{x})}{\partial \mathbf{x}} = \mathrm{A}^T \quad (1)$$

   solution:

   Lets set $\mathbf{y} = \mathrm{A}\mathbf{x}$, by definition (appendix D), we have:
   $$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \left[\begin{array}{cccc} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{array}\right]$$

   First lets compute $\mathbf{y} = \mathrm{A}\mathbf{x}$:
   $$\begin{array}{rcl} \mathbf{y} & = & \left[\begin{array}{cccc} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{array}\right] \times \left(\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right) \\ & = & \left(\begin{array}{c} \sum_{i=1}^n a_{1,i} x_i \\ \sum_{i=1}^n a_{2,i} x_i \\ \vdots \\ \sum_{i=1}^n a_{m,i} x_i \end{array}\right) \end{array}$$

   So we can compute all the derivatives:
   $$\frac{\partial y_1}{\partial x_1} = \frac{\partial(\sum_{i=1}^n a_{1,i} x_i)}{\partial x_1} = a_{1,1}$$
   and
   $$\frac{\partial y_1}{\partial x_2} = \frac{\partial(\sum_{i=1}^n a_{1,i} x_i)}{\partial x_2} = a_{1,2}$$
   and
   $$\frac{\partial y_1}{\partial x_3} = \frac{\partial(\sum_{i=1}^n a_{1,i} x_i)}{\partial x_3} = a_{1,3}$$
   and
   $$\cdots \text{etc.}$$
   and more generally
   $$\frac{\partial y_k}{\partial x_j} = \frac{\partial(\sum_{i=1}^n a_{k,i} x_i)}{\partial x_j} = a_{k,j}$$
   therefore
   $$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \left[\begin{array}{cccc} \frac{\partial y_1}{\partial x_1} = a_{1,1} & \frac{\partial y_2}{\partial x_1} = a_{2,1} & \cdots & \frac{\partial y_m}{\partial x_1} = a_{m,1} \\ \frac{\partial y_1}{\partial x_2} = a_{1,2} & \frac{\partial y_2}{\partial x_2} = a_{2,2} & \cdots & \frac{\partial y_m}{\partial x_2} = a_{m,2} \\ \\ \frac{\partial y_1}{\partial x_n} = a_{1,n} & \frac{\partial y_2}{\partial x_n} = a_{2,n} & \cdots & \frac{\partial y_m}{\partial x_n} = a_{m,n} \end{array}\right] = \mathrm{A}^T$$

(b) Show that

$$\frac{\partial(\mathbf{x}^T \mathrm{A})}{\partial \mathbf{x}} = \mathrm{A} \tag{2}$$

(c) Show that

$$\frac{\partial(\mathbf{x}^T \mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{x} \tag{3}$$

(d) (Assignment) Show that

$$\frac{\partial(\mathbf{x}^T \mathrm{A}\mathbf{x})}{\partial \mathbf{x}} = \mathrm{A}\mathbf{x} + \mathrm{A}^T\mathbf{x} \tag{4}$$

2. **Least square estimation.**

(a) Using equations (1), (2),(3) and (4), find the least square estimate $\mathbf{x}_{LS}$ such that:

$$\mathbf{x}_{LS} = \arg\min_{\mathbf{x}} \left\{ \mathcal{J}(\mathbf{x}) = \|\mathrm{A}\mathbf{x} - \mathbf{y}\|^2 \right\}$$

we did already the computation in the lecture and you should find:

$$\mathbf{x}_{LS} = (\mathrm{A}^T\mathrm{A})^{-1}\mathrm{A}^T\mathbf{y}$$

(b) Note that the previous problem can be rewritten as finding the least square estimate $\mathbf{x}_{LS}$ such that:

$$\mathbf{x}_{LS} = \arg\min_{\mathbf{x}} \left\{ \mathcal{J}(\mathbf{x}) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{j,i}x_i - y_j \right)^2 \right\} \tag{5}$$

(Assignment) Show that we have the same result as before for $\mathbf{x}_{LS}$ using this formulation (5).

3. **Principal component analysis.**

(a) Show that

$$\begin{aligned} \|\mathbf{v}\mathbf{v}^T\mathbf{x}_i - \mathbf{v}\mathbf{v}^T\mathbf{x}_j\|^2 &= (\mathbf{v}\mathbf{v}^T\mathbf{x}_i - \mathbf{v}\mathbf{v}^T\mathbf{x}_j)^T(\mathbf{v}\mathbf{v}^T\mathbf{x}_i - \mathbf{v}\mathbf{v}^T\mathbf{x}_j) \\ &= \mathbf{v}^T(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T\mathbf{v} \end{aligned}$$

(b) Show that

$$\begin{aligned} \mathcal{J}(\mathbf{v}) &= \tfrac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{v}^T(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T\mathbf{v} \\ &= \mathbf{v}^T\mathrm{C}\mathbf{v} \end{aligned}$$

(c) Using equations (1), (2),(3) and (4), find the solution of:

$$\mathbf{v}_{pca} = \arg\max_{\mathbf{v}} \left\{ \mathbf{v}^T\mathrm{C}\mathbf{v} - \lambda(\mathbf{v}^T\mathbf{v} - 1) \right\}$$