University of California, Los Angeles
Department of Statistics

Statistics 100C **Instructor: Nicolas Christou**

Matrix and vector differentiation

Let

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{pmatrix}$$

be a p-dimensional vector and let $f(\boldsymbol{\theta})$ be a function of $\boldsymbol{\theta}$. When the derivative of $f(\boldsymbol{\theta})$ is taken with respect to the vector $\boldsymbol{\theta}$ we mean that the partial derivative of $f(\boldsymbol{\theta})$ is taken with respect to each element of $\boldsymbol{\theta}$, i.e.

$$\frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial f(\boldsymbol{\theta})}{\partial \theta_1} \\ \frac{\partial f(\boldsymbol{\theta})}{\partial \theta_2} \\ \vdots \\ \frac{\partial f(\boldsymbol{\theta})}{\partial \theta_p} \end{pmatrix}$$

We will present now two important results of matrix differentiation.

1. Let $\boldsymbol{\theta}$ as defined above and $\mathbf{c}' = (c_1, c_2, \ldots, c_p)$. If $f(\boldsymbol{\theta}) = \mathbf{c}'\boldsymbol{\theta}$ it follows that

$$\frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{c}.$$

Proof

$$\frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} \frac{\partial f(\boldsymbol{\theta})}{\partial \theta_1} \\ \frac{\partial f(\boldsymbol{\theta})}{\partial \theta_2} \\ \vdots \\ \frac{\partial f(\boldsymbol{\theta})}{\partial \theta_p} \end{pmatrix} = \begin{pmatrix} \frac{\partial \mathbf{c}'\boldsymbol{\theta}}{\partial \theta_1} \\ \frac{\partial \mathbf{c}'\boldsymbol{\theta}}{\partial \theta_2} \\ \vdots \\ \frac{\partial \mathbf{c}'\boldsymbol{\theta})}{\partial \theta_p} \end{pmatrix} = \begin{pmatrix} \frac{\partial (c_1\theta_1 + \ldots + c_p\theta_p)}{\partial \theta_1} \\ \frac{\partial (c_1\theta_1 + \ldots + c_p\theta_p)}{\partial \theta_2} \\ \vdots \\ \frac{\partial (c_1\theta_1 + \ldots + c_p\theta_p)}{\partial \theta_p} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{pmatrix} = \mathbf{c}.$$

2. Let $\mathbf{A}$ be a $p \times p$ symmetric matrix and let $\boldsymbol{\theta}$ as define above. Define now the quadratic expression $f(\boldsymbol{\theta}) = \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta}$. It follows that

$$\frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 2\mathbf{A}\boldsymbol{\theta}.$$

Proof
Let

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1p} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & a_{p3} & \ldots & a_{pp} \end{pmatrix} = \begin{pmatrix} \mathbf{a_1'} \\ \mathbf{a_2'} \\ \vdots \\ \vdots \\ \mathbf{a_p'} \end{pmatrix}.$$

We can write $f(\boldsymbol{\theta})$ as: $f(\boldsymbol{\theta}) = \boldsymbol{\theta}'\mathbf{A}\boldsymbol{\theta} = \sum_{i=1}^{p} \theta_i^2 a_{ii} + 2\sum_{i=1}^{p-1}\sum_{j>i}^{p} \theta_i\theta_j a_{ij}$.
Take the derivative of $f(\boldsymbol{\theta})$ with respect to θ_1: $\frac{\partial f(\boldsymbol{\theta})}{\partial \theta_1} = 2a_{11}\theta_1 + 2\sum_{j\neq 1}^{p} a_{1j}\theta_j = 2\sum_{j=1}^{p} a_{1j}\theta_j = 2\mathbf{a_1'}\boldsymbol{\theta}$.

Take the derivative of $f(\boldsymbol{\theta})$ with respect to θ_2: $\frac{\partial f(\boldsymbol{\theta})}{\partial \theta_2} = 2a_{22}\theta_2 + 2\sum_{j\neq 2}^{p} a_{2j}\theta_j = 2\sum_{j=1}^{p} a_{2j}\theta_j = 2\mathbf{a_2'}\boldsymbol{\theta}$.

$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$

Take the derivative of $f(\boldsymbol{\theta})$ with respect to θ_p: $\frac{\partial f(\boldsymbol{\theta})}{\partial \theta_p} = 2a_{pp}\theta_p + 2\sum_{j\neq p}^{p} a_{pj}\theta_j = 2\sum_{j=1}^{p} a_{pj}\theta_j = 2\mathbf{a_p'}\boldsymbol{\theta}$.

Therefore,

$$\frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{pmatrix} 2\mathbf{a_1'}\boldsymbol{\theta} \\ 2\mathbf{a_2'}\boldsymbol{\theta} \\ \vdots \\ 2\mathbf{a_p'}\boldsymbol{\theta} \end{pmatrix} = 2\begin{pmatrix} \mathbf{a_1'} \\ \mathbf{a_2'} \\ \vdots \\ \mathbf{a_p'} \end{pmatrix}\boldsymbol{\theta} = 2\mathbf{A}\boldsymbol{\theta}.$$