Partial derivatives of $||Ax||^2$ and $2b^TAx$

Kevin Jia, 28 July 2020

1 Proof that the partial derivatives of $||Ax||^2$ with respect to x_l equals $(2A^TAx)_l$ (element l of $2A^TAx$)

1.1: Definitions of matrix A and vectors x and b:

x, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times 2}$. First column of A is all 1's.

1.2: Derivation of partial derivatives of $||Ax||^2$ with respect to x_l:

$f(x) = ||Ax||^2$
$= (Ax)^T Ax$
$= \sum_{i=1}^{m}(Ax)^T Ax$
$= \sum_{i=1}^{m}(\sum_{j=1}^{m} a_{ij}x_j)(\sum_{k=1}^{m} a_{ik}x_k)^{*1}$

^{*1}(The reason that the sigma form of $(Ax)^T$ is the same as the sigma form for Ax (except for swapped variables – j for $(Ax)^T$ and k for Ax), is because sigma sums do not care whether a vector is a row or column one – the values inside a column vector and its respective row vector transpose are the same $((a^T)_i = a_i)$, hence the virtually same sigma form for Ax and it's transposed sibling $(Ax)^T$)

$\therefore \frac{\partial f(x)}{\partial x_l} = \frac{\partial}{\partial x_l}(\sum_{i=1}^{m}\sum_{j=1}^{m} a_{ij}x_j \sum_{k=1}^{m} a_{ik}x_k)$
$= \frac{\partial}{\partial x_l}(\sum_{i=1}^{m}(a_{il}x_l \sum_{k=1,k\neq l}^{m} a_{ik}x_k + a_{il}^2x_l^2 + a_{il}x_l \sum_{k=1,k\neq l}^{m} a_{ik}x_k))^{*2}$

^{*2}(Think of $(a_{11}x_1+a_{12}x_2+a_{13}x_3+...a_{1m}x_m)(a_{11}x_1+a_{12}x_2+a_{13}x_3+...a_{1m}x_m)$, and what terms have x_l in them, as x_l is the term we are taking the partial derivative of $||Ax||^2$ with respect to – all other terms without x_l can be neglected because they will be treated as constants and constants always differentiate to 0)

$= \sum_{i=1}^{m}(a_{il}\sum_{k=1,k\neq l}^{m} a_{ik}x_k + 2a_{il}^2x_l + \sum_{k=1,k\neq l}^{m} a_{ik}x_k)$
$= 2\sum_{i=1}^{m}(a_{il}\sum_{k=1,k\neq l}^{m} a_{ik}x_k + a_{il}^2x_l)^{*3}$

^{*3}(The $a_{il}^2x_l$ is equal to $a_{il}\sum_{k=1}^{m} a_{ik}x_k$ if we allow $k = 1$, so we can combine the formerly mentioned term $a_{il}^2x_l$ with $a_{il}\sum_{k=1,k\neq l}^{m} a_{ik}x_k$ to produce

$a_{il}\sum_{k=1}^{m} a_{ik}x_k)$
$= 2\sum_{i=1}^{m} a_{il}\sum_{k=1}^{m} a_{ik}x_k$
$= 2\sum_{i=1}^{m}\sum_{k=1}^{m} a_{il}a_{ik}x_k$
$\therefore \frac{\partial f(x)}{\partial x_l} = 2\sum_{i=1}^{m}\sum_{k=1}^{m} a_{il}a_{ik}x_k{}^{*4}$
^{*4}(a_{il} does not depend on or change with changes to k, so we can move a_{il} within $2\sum_{i=1}^{m}\sum_{k=1}^{m} a_{il}a_{ik}x_k$)

1.3: Calculation of the expression defining $2A^TAx_l$ (element l of A^TAx):

$Ax_i = \sum_{k=1}^{m} a_{ik}x_k$

$2A^TAx_l = 2\sum_{i=1}^{m}(A^T)_{li}(Ax)_i$
$= 2\sum_{i=1}^{m} a_{il}(Ax)_i{}^{*5}$

^{*5}(Element a_{ij} of matrix A would have element a_{ji} (also of matrix A) in its position for transposed matrix A^T. This means $(A^T)_{ji} = a_{ij}$, hence the switch of variables i and l from the previous expression)
$= 2\sum_{i=1}^{m} a_{il}\sum_{k=1}^{m} a_{ik}x_k$
$= 2\sum_{i=1}^{m}\sum_{k=1}^{m} a_{il}a_{ik}x_k{}^{*6}$

^{*6}(See ^{*4})
$= \frac{\partial f(x)}{\partial x_l}$
$\therefore 2A^TAx_l = \frac{\partial f(x)}{\partial x_l}$
$QED.$

2 Proof that the partial derivatives of $2b^TAx$ with respect to x_k equals $(2b^TA)_k$ (element k of $2b^TA$)

2.1: Derivation of partial derivatives of $2b^TAx$ with respect to x_k:

$Ax_i = \sum_{j=1}^{m} a_{ij}x_j$

$g(x) = 2b^TAx$
$= 2\sum_{i=1}^{m}(b^T)_i(Ax)_i$
$= \sum_{i=1}^{m} b_i(Ax)_i{}^{*7}$

^{*7}(See ^{*1} – the part about indifferent sigma sums and column vectors and their

transposes)
$= \sum_{i=1}^{m} b_i \sum_{j=1}^{m} a_{ij} x_j$
$= \sum_{i=1}^{m} \sum_{j=1}^{m} b_i a_{ij} {x_j}^{*8}$

^{*8}(See ^{*4})
$= \sum_{i=1}^{m} (\sum_{j=1, j \neq k}^{m} b_i a_{ij} x_j + b_i a_{ik} x_k)^{*9}$

^{*9}(See ^{*2}, but replace x_l with x_k and replace $||Ax||^2$ with $2b^T Ax$)
$\therefore \frac{\partial g(x)}{\partial x_k} = \sum_{i=1}^{m} b_i a_{ik}$

2.2: Calculation of expression defining $(2b^T A)_k$ (element k of $2b^T A$):
$2b^T A_k = \sum_{i=1}^{m} (b^T)_i A_{ik}$
$= \sum_{i=1}^{m} b_i {A_{ik}}^{*10}$

^{*10}(See ^{*1} – the part about indifferent sigma sums and column vectors and their transposes)
$= \frac{\partial g(x)}{x_k}$
$\therefore 2b^T A_k = \frac{\partial g(x)}{x_k}$
$QED.$