Comments on the Quadratic Form

Nathan Marianovsky

Definition 1 (Quadratic Form)**.** Formally a quadratic form is defined as:

$$f(\overrightarrow{x}) = \frac{1}{2}\overrightarrow{x}^T A \overrightarrow{x} - \overrightarrow{b}^T x + c$$

where:

$$c \in \mathbb{R}, \quad \overrightarrow{x}, \overrightarrow{b} \in \mathbb{R}^n, \quad \text{and} \quad A \in \mathbb{R}^{n^2}$$

Definition 2 (Del Operator)**.** The *del operator* is formally defined as:

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

where it is known to be a linear operator:

$$\begin{aligned}
\nabla(\alpha f(\overrightarrow{x}) + \beta g(\overrightarrow{x})) &= \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} (\alpha f(\overrightarrow{x}) + \beta g(\overrightarrow{x})) \\
&= \begin{pmatrix} \frac{\partial}{\partial x_1}(\alpha f(\overrightarrow{x}) + \beta g(\overrightarrow{x})) \\ \frac{\partial}{\partial x_2}(\alpha f(\overrightarrow{x}) + \beta g(\overrightarrow{x})) \\ \frac{\partial}{\partial x_3}(\alpha f(\overrightarrow{x}) + \beta g(\overrightarrow{x})) \\ \vdots \\ \frac{\partial}{\partial x_n}(\alpha f(\overrightarrow{x}) + \beta g(\overrightarrow{x})) \end{pmatrix} \\
&= \begin{pmatrix} \alpha \frac{\partial}{\partial x_1} f(\overrightarrow{x}) \\ \alpha \frac{\partial}{\partial x_2} f(\overrightarrow{x}) \\ \alpha \frac{\partial}{\partial x_3} f(\overrightarrow{x}) \\ \vdots \\ \alpha \frac{\partial}{\partial x_n} f(\overrightarrow{x}) \end{pmatrix} + \begin{pmatrix} \beta \frac{\partial}{\partial x_1} g(\overrightarrow{x}) \\ \beta \frac{\partial}{\partial x_2} g(\overrightarrow{x}) \\ \beta \frac{\partial}{\partial x_3} g(\overrightarrow{x}) \\ \vdots \\ \beta \frac{\partial}{\partial x_n} g(\overrightarrow{x}) \end{pmatrix} \\
&= \alpha \nabla f(\overrightarrow{x}) + \beta \nabla g(\overrightarrow{x})
\end{aligned}$$

Definition 3 (Gradient)**.** Given a scalar function that has n parameters such as $f(\overrightarrow{x})$ above, then the *derivative* or *gradient* is formally defined as:

$$f'(\overrightarrow{x}) = \nabla f(\overrightarrow{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} f(\overrightarrow{x}) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Proposition 1. The gradient of the quadratic form is given as:

$$f'(\overrightarrow{x}) = \frac{1}{2}A^T\overrightarrow{x} + \frac{1}{2}A\overrightarrow{x} - \overrightarrow{b}$$

and if A is symmetric:

$$f'(\overrightarrow{x}) = A\overrightarrow{x} - \overrightarrow{b}$$

Proof. To prove the proposition we begin by defining:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

$$\overrightarrow{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \overrightarrow{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

Now using this we have:

$$A\overrightarrow{x} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum a_{1i}x_i \\ \sum a_{2i}x_i \\ \vdots \\ \sum a_{ni}x_i \end{pmatrix}$$

$$A^T\overrightarrow{x} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum a_{i1}x_i \\ \sum a_{i2}x_i \\ \vdots \\ \sum a_{in}x_i \end{pmatrix}$$

where the sums are iterated from $i = 1$ to n to account for all the components. Now the next step is:

$$\begin{aligned} \overrightarrow{x}^T A\overrightarrow{x} &= \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} \sum a_{1i}x_i \\ \sum a_{2i}x_i \\ \vdots \\ \sum a_{ni}x_i \end{pmatrix} \\ &= x_1 \sum a_{1i}x_i + x_2 \sum a_{2i}x_i + \dots + x_n \sum a_{ni}x_i \\ &= \sum x_j \sum a_{ji}x_i \\ &= \sum \sum a_{ji}x_i x_j \end{aligned}$$

With this we have a nice consolidated form for that annoying matrix multiplication. Now we want to find the partial derivatives. To make it easy on the brain I will find the partial derivative with respect to x_1 and then generalize it. So we have:

$$\begin{aligned} \frac{\partial(\overrightarrow{x}^T A\overrightarrow{x})}{\partial x_1} &= \frac{\partial}{\partial x_1}\left[x_1 \sum a_{1i}x_i + x_2 \sum a_{2i}x_i + \dots + x_n \sum a_{ni}x_i\right] \\ &= (2a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) + (a_{21}x_2 + a_{31}x_3 + \dots + a_{n1}x_n) \\ &= \sum a_{1i}x_i + \sum a_{i1}x_i \end{aligned}$$

Since $k = 1$ is a specific case, in general we have:

$$\frac{\partial(\overrightarrow{x}^T A\overrightarrow{x})}{\partial x_k} = \sum a_{ki}x_i + \sum a_{ik}x_i$$

Now all of the hard work is done. Phew! All that is left is to simplify our expression. We begin by writing down the gradient:

$$
\begin{aligned}
\nabla(\overrightarrow{x}^T A \overrightarrow{x}) &= \begin{pmatrix} \sum a_{1i}x_i + \sum a_{i1}x_i \\ \sum a_{2i}x_i + \sum a_{i2}x_i \\ \vdots \\ \sum a_{ni}x_i + \sum a_{in}x_i \end{pmatrix} \\
&= \begin{pmatrix} \sum a_{1i}x_i \\ \sum a_{2i}x_i \\ \vdots \\ \sum a_{ni}x_i \end{pmatrix} + \begin{pmatrix} \sum a_{i1}x_i \\ \sum a_{i2}x_i \\ \vdots \\ \sum a_{in}x_i \end{pmatrix} \\
&= A\overrightarrow{x} + A^T\overrightarrow{x}
\end{aligned}
$$

Lastly before jumping into the quadratic form we define:

$$
\begin{aligned}
\overrightarrow{b}^T\overrightarrow{x} &= \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\
&= \sum b_i x_i
\end{aligned}
$$

where the gradient is going to be:

$$
\begin{aligned}
\nabla(\overrightarrow{b}^T\overrightarrow{x}) &= \nabla\Big(\sum b_i x_i\Big) \\
&= \begin{pmatrix} \frac{\partial}{\partial x_1}\sum b_i x_i \\ \frac{\partial}{\partial x_2}\sum b_i x_i \\ \vdots \\ \frac{\partial}{\partial x_n}\sum b_i x_i \end{pmatrix} \\
&= \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \\
&= \overrightarrow{b}
\end{aligned}
$$

Now using our results, the gradient of the quadratic form is going to be:

$$
\begin{aligned}
f'(\overrightarrow{x}) &= \nabla f(\overrightarrow{x}) \\
&= \frac{1}{2}\nabla(\overrightarrow{x}^T A \overrightarrow{x}) - \nabla(\overrightarrow{b}^T\overrightarrow{x}) + \nabla(c) \\
&= \frac{1}{2}(A\overrightarrow{x} + A^T\overrightarrow{x}) - (\overrightarrow{b}) + (\overrightarrow{0}) \\
&= \frac{1}{2}A\overrightarrow{x} + \frac{1}{2}A^T\overrightarrow{x} - \overrightarrow{b}
\end{aligned}
$$

Lastly if A is symmetric that means $A = A^T$ which will give:

$$
f'(\overrightarrow{x}) = A\overrightarrow{x} - \overrightarrow{b}
$$

■