5.2.4 Analyse matricielle

❒ **Gradient** – Soit $f:\mathbb{R}^{m\times n}\rightarrow\mathbb{R}$ une fonction et $A\in\mathbb{R}^{m\times n}$ une matrice. Le gradient de f par rapport à A est une matrice de taille $m\times n$, notée $\nabla_A f(A)$, telle que :

$$\boxed{\Big(\nabla_A f(A)\Big)_{i,j}=\frac{\partial f(A)}{\partial A_{i,j}}}$$

Remarque : le gradient de f est seulement défini lorsque f est une fonction donnant un scalaire.

❒ **Hessienne** – Soit $f:\mathbb{R}^{n}\rightarrow\mathbb{R}$ une fonction et $x\in\mathbb{R}^{n}$ un vecteur. La hessienne de f par rapport à x est une matrice symmetrique $n\times n$, notée $\nabla_x^2 f(x)$, telle que :

$$\boxed{\Big(\nabla_x^2 f(x)\Big)_{i,j}=\frac{\partial^2 f(x)}{\partial x_i\partial x_j}}$$

Remarque : la hessienne de f est seulement définie lorsque f est une fonction qui donne un scalaire.

❒ **Opérations de gradient** – Pour des matrices A,B,C, les propriétés de gradient suivants sont bons à savoir :

$$\boxed{\nabla_A\textrm{tr}(AB)=B^T}\qquad\boxed{\nabla_{A^T}f(A)=(\nabla_Af(A))^T}$$

$$\boxed{\nabla_A\textrm{tr}(ABA^TC)=CAB+C^TAB^T}\qquad\boxed{\nabla_A|A|=|A|(A^{-1})^T}$$