一般方法：Taylor展开：

$f(x+h) = f(x) + (\nabla f(x))^T \cdot h + \frac{1}{2} h^T \nabla^2 f(x) h + \cdots$

函数 x+h 在 x 处泰勒展开的公式比较

向量、矩阵求导速记表：

- $(a^T x)' = a$
- $(x^T a)' = a$
- $(x^T x)' = 2x$

  $(x^T x)'' = 2$
- $((x+a)^T(x+a))' = 2(x+a)$

  $((x+a)^T(x+a))'' = 2I$
- $(x^T A x)' = Ax + A^T x$

  $(x^T A x)'' = A + A^T$

找这组合到参量

模长的含到向量

类二次型矩阵.

规律：向量一阶梯是到向量.
不会做就用Taylor

复合函数：　$f(x) = \frac{1}{2}\|Ax-b\|_2^2 = \frac{1}{2}(Ax-b)^T(Ax-b)$

By taylor：
$$f(x+h) = \frac{1}{2}(Ax+Ah-b)^T(Ax+Ah-b)$$
$$= \frac{1}{2}(Ax-b)^T(Ax-b) + \frac{1}{2}(Ah)^T(Ax-b) + \frac{1}{2}(Ax-b)^T Ah + \frac{1}{2}(Ah)^T Ah .$$
$$= \underbrace{\frac{1}{2}(Ax-b)^T(Ax-b)}_{f(x)} + \underbrace{(A^T(Ax-b))^T}_{\nabla f(x)} h + \frac{1}{2} h^T \underbrace{A^T A}_{\nabla^2 f(x)} h$$

By 速记. ：
$$\nabla f(x) = A^T(Ax-b) + \mu x$$
$$\nabla^2 f(x) = A^T A + \mu I .$$
$$\succeq \mu I \succ 0 .$$

By 复合 ：
令 $h(y) = \frac{1}{2}\|y\|_2^2 = \frac{1}{2} y^T y .$

有 $f(x) = h(Ax-b)$
$$\nabla f(x) = A^T \nabla h(Ax-b)$$
$$\nabla^2 f(x) = A^T \nabla^2 h(Ax-b) A$$
其中 $\nabla^2 h(y) = I$

所以：$\nabla^2 f(x) = A^T A .$

① 复合函数求导 写在前面：求导的函数放前面！
② 代入求导 ≠ 求导代入！ $h'(Ax-b) \neq \nabla h(Ax-b)$

---

1. Let $G \in \mathbb{R}^{n\times n}$ be a symmetric matrix, $c \in \mathbb{R}^n$, and define

$$f(x) := \frac{1}{2} x^T G x + c^T x$$

Using Taylor's theorem, argue that $\nabla f(x) = Gx + c$.

$f(x) = \frac{1}{2} x^T G x + c^T x$

$f(x+h) = \frac{1}{2}(x+h)^T G (x+h) + c^T(x+h)$
$$= \cdots + \frac{1}{2}(h^T G x + x^T G h) + c^T h + \cdots \quad I$$
$$= \cdots + \frac{1}{2}(x^T G^T h + x^T G h) + c^T h + \cdots$$
$$= \cdots + (x^T G + c^T) h + \cdots$$
$$= \cdots + \underbrace{(Gx + c)^T}_{\nabla f(x)} \cdot h$$

I 泰勒展开法求矩阵导数.
　II 是利用 f(x+h) 的 T.E. 通过对应项求.