证明. 先证必要性. 设 f 是凸函数，则对于任意的 $x,y \in \mathbf{dom}\, f$ 以及 $t \in (0,1)$，有

$$tf(y) + (1-t)f(x) \geqslant f(x + t(y-x)).$$

将上式移项，两边同时除以 t，注意 $t>0$，则

$$f(y) - f(x) \geqslant \frac{f(x+t(y-x)) - f(x)}{t}.$$

令 $t \to 0$，由极限保号性可得

$$f(y) - f(x) \geqslant \lim_{t\to 0} \frac{f(x+t(y-x)) - f(x)}{t} = \nabla f(x)^{\mathrm{T}}(y-x).$$

这里最后一个等式成立是由于方向导数的性质.

再证充分性. 对任意的 $x,y \in \mathbf{dom}\, f$ 以及任意的 $t \in (0,1)$，定义 $z = tx + (1-t)y$，应用两次一阶条件我们有

$$\begin{aligned} f(x) &\geqslant f(z) + \nabla f(z)^{\mathrm{T}}(x-z), \\ f(y) &\geqslant f(z) + \nabla f(z)^{\mathrm{T}}(y-z). \end{aligned}$$

将上述第一个不等式两边同时乘 t，第二个不等式两边同时乘 $1-t$，相加得

$$tf(x) + (1-t)f(y) \geqslant f(z) + 0.$$

这正是凸函数的定义，因此充分性成立. □

定理 2.9 说明可微凸函数 f 的图形始终在其任一点处切线的上方，见图 2.12. 因此，用可微凸函数 f 在任意一点处的一阶近似可以得到 f 的一个全局下界. 另一个常用的一阶条件是梯度单调性.

定理 2.10 (梯度单调性) 设 f 为可微函数，则 f 为凸函数当且仅当 $\mathbf{dom}\, f$ 为凸集且 ∇f 为**单调映射**，即

$$(\nabla f(x) - \nabla f(y))^{\mathrm{T}}(x-y) \geqslant 0, \quad \forall\, x,y \in \mathbf{dom}\, f.$$

证明. 先证必要性. 若 f 可微且为凸函数，根据一阶条件，我们有

$$\begin{aligned} f(y) &\geqslant f(x) + \nabla f(x)^{\mathrm{T}}(y-x), \\ f(x) &\geqslant f(y) + \nabla f(y)^{\mathrm{T}}(x-y). \end{aligned}$$

将两式不等号左右两边相加即可得到结论.

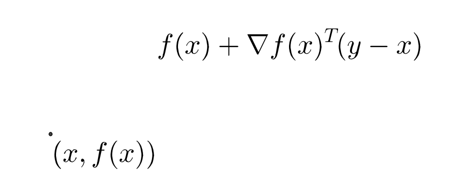

图 2.12 凸函数的全局下界

再证充分性. 若 ∇f 为单调映射，构造一元辅助函数

$$g(t)=f(x+t(y-x)),\quad g'(t)=\nabla f(x+t(y-x))^{\mathrm{T}}(y-x)$$

由 ∇f 的单调性可知 $g'(t)\geqslant g'(0),\forall\, t\geqslant 0$. 因此

$$\begin{aligned} f(y)=g(1)&=g(0)+\int_0^1 g'(t)\mathrm{d}t \\ &\geqslant g(0)+g'(0)=f(x)+\nabla f(x)^{\mathrm{T}}(y-x). \end{aligned}$$ □

和凸函数类似，严格凸函数和强凸函数都有对应的单调性.

推论 2.2 设 f 为可微函数，且 $\mathbf{dom}\, f$ 是凸集，则

(1) f 是严格凸函数当且仅当

$$(\nabla f(x)-\nabla f(y))^{\mathrm{T}}(x-y)>0,\quad \forall\, x,y\in\mathbf{dom}\, f;$$

(2) f 是 m-强凸函数当且仅当

$$(\nabla f(x)-\nabla f(y))^{\mathrm{T}}(x-y)\geqslant m\|x-y\|^2,\quad \forall\, x,y\in\mathbf{dom}\, f.$$

进一步地，如果函数二阶连续可微，我们可以得到下面的二阶条件:

定理 2.11 (二阶条件) 设 f 为定义在凸集上的二阶连续可微函数，则 f 是凸函数当且仅当

$$\nabla^2 f(x)\succeq 0,\quad \forall x\in\mathbf{dom}\, f.$$

如果 $\nabla^2 f(x)\succ 0,\forall x\in\mathbf{dom}\, f$，则 f 是严格凸函数.

证明． 先证必要性．反设 $f(x)$ 在点 x 处的海瑟矩阵 $\nabla^2 f(x) \not\succeq 0$，即存在非零向量 $v \in \mathbb{R}^n$ 使得 $v^{\mathrm{T}}\nabla^2 f(x)v < 0$．根据佩亚诺（Peano）余项的泰勒展开，

$$f(x+tv) = f(x) + t\nabla f(x)^{\mathrm{T}}v + \frac{t^2}{2}v^{\mathrm{T}}\nabla^2 f(x)v + o(t^2).$$

移项后等式两边同时除以 t^2，

$$\frac{f(x+tv) - f(x) - t\nabla f(x)^{\mathrm{T}}v}{t^2} = \frac{1}{2}v^{\mathrm{T}}\nabla^2 f(x)v + o(1).$$

当 t 充分小时，

$$\frac{f(x+tv) - f(x) - t\nabla f(x)^{\mathrm{T}}v}{t^2} < 0,$$

这显然和一阶条件（定理 2.9）矛盾，因此必有 $\nabla^2 f(x) \succeq 0$ 成立．

再证充分性．设 $f(x)$ 满足二阶条件 $\nabla^2 f(x) \succeq 0$，对任意 $x, y \in \mathbf{dom}\, f$，根据泰勒展开（定理 2.1），

$$f(y) = f(x) + \nabla f(x)^{\mathrm{T}}(y-x) + \frac{1}{2}(y-x)^{\mathrm{T}}\nabla^2 f(x+t(y-x))(y-x),$$

其中 $t \in (0,1)$ 是和 x, y 有关的常数．由半正定性可知对任意 $x, y \in \mathbf{dom}\, f$ 有

$$f(y) \geqslant f(x) + \nabla f(x)^{\mathrm{T}}(y-x).$$

这是凸函数判定的一阶条件，由定理 2.9 知 f 为凸函数．进一步，若 $\nabla^2 f(x) \succ 0$，上式中不等号严格成立（$x \neq y$）．利用定理 2.9 的充分性的证明过程可得 $f(x)$ 为严格凸函数． □

当函数二阶连续可微时，利用二阶条件判断凸性通常更为方便．下面给出两个用二阶条件判断凸性的例子．

例 2.6

(1) 考虑二次函数 $f(x) = \frac{1}{2}x^{\mathrm{T}}Px + q^{\mathrm{T}}x + r$ $(P \in \mathcal{S}^n)$，容易计算出其梯度与海瑟矩阵分别为

$$\nabla f(x) = Px + q, \quad \nabla^2 f(x) = P.$$

那么，f 是凸函数当且仅当 $P \succeq 0$．

(2) 考虑最小二乘函数 $f(x) = \frac{1}{2}\|Ax - b\|_2^2$，其梯度与海瑟矩阵分别为

$$\nabla f(x) = A^{\mathrm{T}}(Ax - b), \quad \nabla^2 f(x) = A^{\mathrm{T}}A.$$

注意到 $A^{\mathrm{T}}A$ 恒为半正定矩阵，因此，对任意的 A，f 都是凸函数．