$$= \lim_{k\to\infty} \underbrace{(t_k - \|\vec{x}_k\|_2)}_{\geq 0} \tag{5.89}$$

$$\geq 0. \tag{5.90}$$

Thus $(\vec{x}, t) \in K$, so K contains its limits and is closed.

We have proved that K is a convex, pointed, solid, and closed cone, so it is proper.

(b) We show that the dual cone of K in $\mathbb{R}^{n+1}$ is K itself. Let $K^\star = \{(\vec{y}, s) \in \mathbb{R}^{n+1} \mid (\vec{y}, s)^\top (\vec{x}, t) \geq 0 \text{ for all } (\vec{x}, t) \in K\}$ be the dual cone of K. We first show that $K^\star \subseteq K$, then show that $K^\star \supseteq K$.

First, to show that $K^\star \subseteq K$, fix $(\vec{y}, s) \in K^\star$. We want to show that $s \geq \|\vec{y}\|_2$, so that $(\vec{y}, s) \in K^\star$. Since $(\vec{0}, 1) \in K$, by definition of $K^\star$ we have

$$0 \leq (\vec{y}, s)^\top (\vec{0}, 1) = s. \tag{5.91}$$

Thus, if $\|\vec{y}\|_2 = 0$ (i.e., if $\vec{y} = \vec{0}$), then we have $s \geq \|\vec{y}\|_2$, so that $(\vec{y}, s) \in K$.

Now suppose that $\vec{y} \neq 0$, so that $\|\vec{y}\|_2 > 0$. Then $(-\vec{y}, \|\vec{y}\|_2) \in K$, so that

$$0 \leq (\vec{y}, s)^\top (-\vec{y}, \|\vec{y}\|_2) = -\|\vec{y}\|_2^2 + s\|\vec{y}\|_2 \tag{5.92}$$

$$\Rightarrow s\|\vec{y}\|_2 \geq \|\vec{y}\|_2^2 \tag{5.93}$$

$$\Rightarrow s \geq \|\vec{y}\|_2. \tag{5.94}$$

Therefore, $(\vec{y}, s) \in K$. Since $(\vec{y}, s) \in K^\star$ were arbitrary, we have $K^\star \subseteq K$.

Now we want to show that $K \subseteq K^\star$. To this end, let $(\vec{y}, s) \in K$. We want to show that $(\vec{y}, s) \in K^\star$, or equivalently, $(\vec{x}, t)^\top (\vec{y}, s) \geq 0$ for all $(\vec{x}, t) \in K$. Indeed,

$$(\vec{x}, t)^\top (\vec{y}, s) = \vec{x}^\top \vec{y} + st \geq -\|\vec{x}\|_2 \|\vec{y}\|_2 - st \geq 0 \tag{5.95}$$

where the first inequality follows by Cauchy-Schwarz inequality, and the second inequality follows from the fact that since $(\vec{x}, t), (\vec{y}, s) \in K$ we have $\|\vec{x}\|_2 \leq t$ and $\|\vec{y}\|_2 \leq s$. This shows that $(\vec{y}, s) \in K^\star$, so that $K \subseteq K^\star$.

We have shown that $K \supseteq K^\star$ and $K \subseteq K^\star$, so $K = K^\star$.

□

Theorem 118

Let $K \subseteq \mathbb{R}^n$ be a non-empty closed convex cone. Then $(K^\star)^\star = K$.

Proof. We want to show that $K \subseteq (K^\star)^\star$ and $K \supseteq (K^\star)^\star$.

First, we want to show that $K \subseteq (K^\star)^\star$. Fix $\vec{x} \in K$. We want to show that $\vec{x} \in (K^\star)^\star$, which means that $\vec{x}^\top \vec{y} \geq 0$ for any $\vec{y} \in K^\star$. For this $\vec{y}$, we have $\vec{y}^\top \vec{z} \geq 0$ for all $\vec{z} \in K$. But this includes $\vec{z} = \vec{x}$, so $\vec{x}^\top \vec{y} = \vec{y}^\top \vec{x} \geq 0$. Thus $\vec{x} \in (K^\star)^\star$. We conclude that $K \subseteq (K^\star)^\star$. (Note that $K \subseteq (K^\star)^\star$ actually holds for any cone K, not just non-empty closed convex cones.)

Next, we want to show that $(K^\star)^\star \subseteq K$. Suppose for the sake of contradiction that $(K^\star)^\star \not\subseteq K$. Then there exists $\vec{y} \in (K^\star)^\star$ such that $\vec{y} \notin K$. Since $\vec{y} \notin K$, we have $\vec{y} \neq \vec{0}$. Because K is a closed convex cone, it is a closed convex set. Since $\{\vec{y}\}$ and K are two disjoint closed convex sets, the Separating Hyperplane Theorem tells us that there exists some nonzero $\vec{w} \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that $\vec{w}^\top \vec{x} > c$ for all $\vec{x} \in K$, and $\vec{w}^\top \vec{y} < c$. Since $\vec{0} \in K$, we have $0 = \vec{w}^\top \vec{0} > c$, i.e., $c < 0$, so $\vec{w}^\top \vec{y} < c < 0$. Since K is a cone, for any $\alpha > 0$ we have

$$\vec{w}^\top (\alpha \vec{x}) > c \tag{5.96}$$

$$\Rightarrow \alpha \vec{w}^\top \vec{x} > c \tag{5.97}$$

$$\Rightarrow \vec{w}^\top \vec{x} > \frac{c}{\alpha}. \tag{5.98}$$

By taking $\alpha \to \infty$ we get that $\vec{w}^\top \vec{x} \geq 0$ for any $\vec{x} \in K$. Thus $\vec{w} \in K^\star$. But we must have $\vec{w}^\top \vec{y} < c < 0$, so $\vec{y} \notin (K^\star)^\star$, a contradiction. Thus $(K^\star)^\star \subseteq K$ and so $(K^\star)^\star = K$. $\square$

The above content is optional/out of scope for this semester, but now we resume the required/in scope content.

5.2 Convex Functions

In this section, we define convex and concave functions, and introduce their properties. At the end, we also deliberate on the special and important example of *affine functions*.

Definition 119 (Convex and Concave Functions)
Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a function. We say that f is *convex* if the domain of f, say Ω, is convex, and, for any $\vec{x}_1, \vec{x}_2 \in \Omega$ and $\theta \in [0, 1]$, we have

$$f(\theta \vec{x}_1 + (1-\theta)\vec{x}_2) \leq \theta f(\vec{x}_1) + (1-\theta) f(\vec{x}_2). \tag{5.99}$$

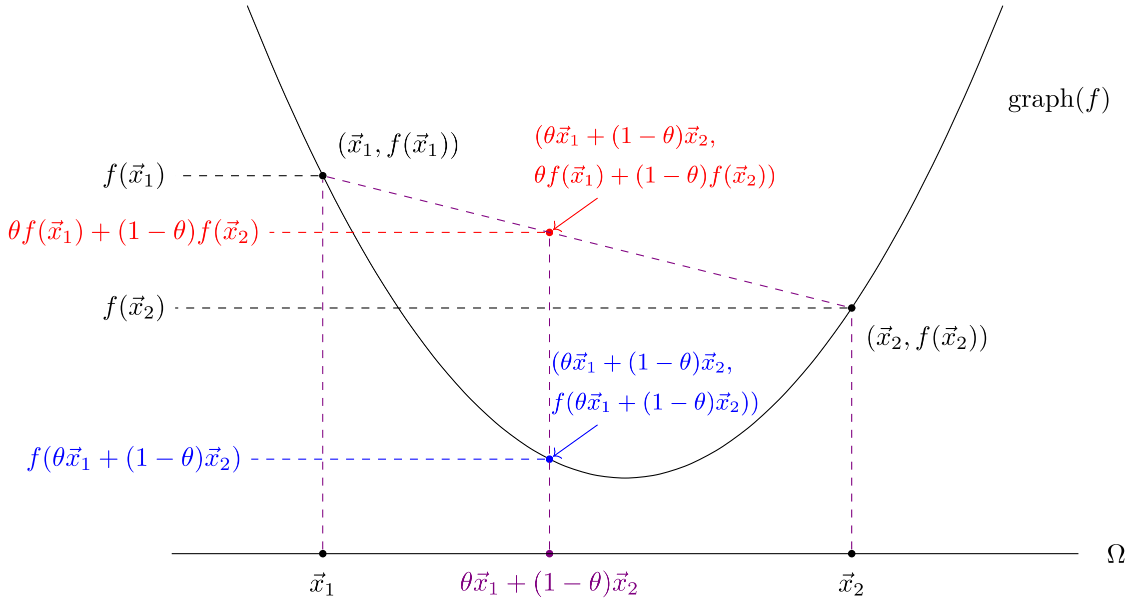
A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is *concave* if $-f$ is convex.

Equation (5.99) is also called *Jensen's inequality* and is equivalent to the following, seemingly more general statement.

Theorem 120 (Jensen's Inequality)
Let Ω be a convex set and let $f\colon \Omega \to \mathbb{R}$ be a convex function. Let $\vec{x}_1, \ldots, \vec{x}_k \in \Omega$, and let $\theta_1, \ldots, \theta_k \in [0, 1]$ such that $\sum_{i=1}^k \theta_i = 1$. Then

$$f\left(\sum_{i=1}^k \theta_i \vec{x}_i\right) \leq \sum_{i=1}^k \theta_i f(\vec{x}_i). \tag{5.100}$$

We can think about this result in terms of a picture.



The prototypical convex function has a "bowl-shaped" graph, and taking a weighted average of two (or any finite number) of points will mean we land in the bowl. In particular, taking a weighted average of any number of function values $f(\vec{x}_1), \ldots, f(\vec{x}_k)$ will always give a larger number than applying the function f to the same weighted averages of the points $\vec{x}_1, \ldots, \vec{x}_k$. Put more simply, if f is convex then the chord joining the points $(\vec{x}_1, f(\vec{x}_1))$ and $(\vec{x}_2, f(\vec{x}_2))$ always lies *above* the graph of f. Similarly, if f is concave then the chord joining the points $(\vec{x}_1, f(\vec{x}_1))$ and $(\vec{x}_2, f(\vec{x}_2))$ always lies *below* the graph of f.

From the picture, it may be intuitively clear that it is hard to construct convex functions f with multiple global minima. We will come back to this idea later.

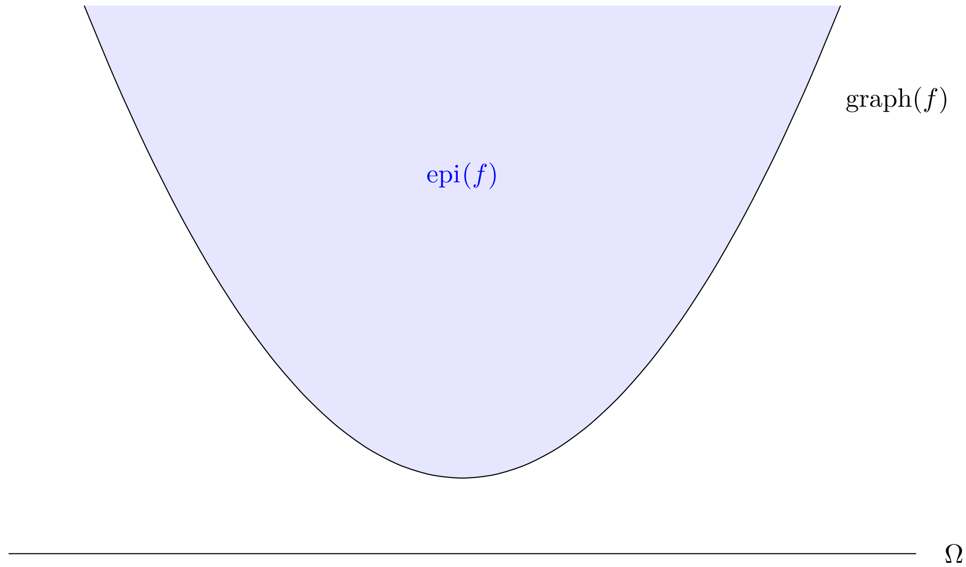
It may be useful to connect the notion of convex function and convex set. For this, we will define the epigraph.

Definition 121 (Epigraph)

Let Ω be a convex set and let $f\colon \Omega \to \mathbb{R}$ be a convex function. The *epigraph* of f, denoted $\operatorname{epi}(f) \subseteq \Omega \times \mathbb{R}$, is defined as

$$\operatorname{epi}(f) = \{(\vec{x}, t) \mid \vec{x} \in \Omega, t \geq f(\vec{x})\}. \tag{5.101}$$

Geometrically, the epigraph is all points in $\Omega \times \mathbb{R}$ that lie above the graph of f.



Proposition 122

Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f \colon \Omega \to \mathbb{R}$ be a function. Then f is a convex function if and only if $\mathrm{epi}(f)$ is a convex set.

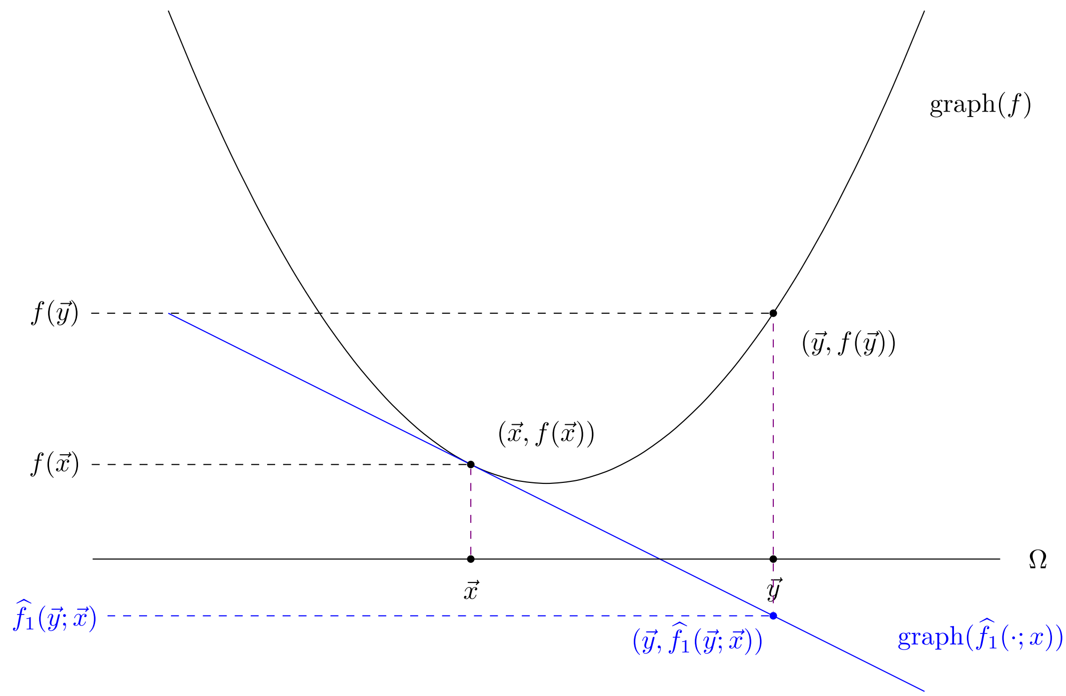
Proof. Left as an exercise. □

Theorem 123 (First-Order Condition for Convexity)

Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f \colon \Omega \to \mathbb{R}$ be a differentiable function. Then f is convex if and only if for all $\vec{x}, \vec{y} \in \Omega$, we have

$$f(\vec{y}) \geq f(\vec{x}) + [\nabla f(\vec{x})]^\top (\vec{y} - \vec{x}). \tag{5.102}$$

Note that this latter term is the first-order Taylor expansion of f around $\vec{x}$ evaluated at $\vec{y}$, i.e., $\widehat{f}_1(\vec{y}; \vec{x}) = f(\vec{x}) + [\nabla f(\vec{x})]^\top (\vec{y} - \vec{x})$. The graph of $\widehat{f}_1(\cdot; \vec{x})$ is the tangent line to the graph of f at the point $(\vec{x}, f(\vec{x}))$. So another characterization of convex functions is that their graphs lie above their tangent lines.



Proof. First suppose f is convex. Then for any $h \in (0, 1)$, we have

$$f(h\vec{y} + (1-h)\vec{x}) \leq hf(\vec{y}) + (1-h)f(\vec{x}) \tag{5.103}$$
$$= hf(\vec{y}) + f(\vec{x}) - hf(\vec{x}) \tag{5.104}$$
$$= f(\vec{x}) + h(f(\vec{y}) - f(\vec{x})) \tag{5.105}$$
$$\implies f(h\vec{y} + (1-h)\vec{x}) - f(\vec{x}) \leq h(f(\vec{y}) - f(\vec{x})) \tag{5.106}$$
$$\implies \frac{f(h\vec{y} + (1-h)\vec{x}) - f(\vec{x})}{h} \leq f(\vec{y}) - f(\vec{x}) \tag{5.107}$$
$$\implies f(\vec{y}) \geq f(\vec{x}) + \frac{f(h\vec{y} + (1-h)\vec{x}) - f(\vec{x})}{h} \tag{5.108}$$
$$= f(\vec{x}) + \frac{f(\vec{x} + h(\vec{y} - \vec{x})) - f(\vec{x})}{h}. \tag{5.109}$$

To summarize, for any $h \in (0, 1)$ we have that

$$f(\vec{y}) \geq f(\vec{x}) + \frac{f(\vec{x} + h(\vec{y} - \vec{x})) - f(\vec{x})}{h}. \tag{5.110}$$

Taking the limit $h \to 0$ on both sides, we get

$$f(\vec{y}) \geq \lim_{h \to 0} \left\{ f(\vec{x}) + \frac{f(\vec{x} + h(\vec{y} - \vec{x})) - f(\vec{x})}{h} \right\} \tag{5.111}$$
$$= f(\vec{x}) + \lim_{h \to 0} \frac{f(\vec{x} + h(\vec{y} - \vec{x})) - f(\vec{x})}{h} \tag{5.112}$$
$$= f(\vec{x}) + [\nabla f(\vec{x})]^\top (\vec{y} - \vec{x}) \tag{5.113}$$

as desired. Here the last equality is because the limit is interpreted as a directional derivative, and it has already been shown that directional derivatives are equal to inner products of the gradient with the direction vector.

For the other direction, let $\theta \in [0, 1]$ and let $\vec{z} = \theta\vec{x} + (1-\theta)\vec{y}$. We have

$$f(\vec{x}) \geq f(\vec{z}) + [\nabla f(\vec{z})]^\top (\vec{x} - \vec{z}) \tag{5.114}$$

$$f(\vec{y}) \geq f(\vec{z}) + [\nabla f(\vec{z})]^\top (\vec{y} - \vec{z}). \tag{5.115}$$

Adding θ times the first equation to $(1-\theta)$ times the second equation, we get

$$\begin{aligned}
\theta f(\vec{x}) + (1-\theta) f(\vec{y}) &\geq \theta f(\vec{z}) + (1-\theta) f(\vec{z}) + \theta[\nabla f(\vec{z})]^\top (\vec{x} - \vec{z}) + (1-\theta)[\nabla f(\vec{z})]^\top (\vec{y} - \vec{z}) && (5.116) \\
&= f(\vec{z}) + [\nabla f(\vec{z})]^\top (\theta \vec{x} + (1-\theta)\vec{y} - \vec{z}) && (5.117) \\
&= f(\vec{z}) + [\nabla f(\vec{z})]^\top (\vec{z} - \vec{z}) && (5.118) \\
&= f(\vec{z}) + [\nabla f(\vec{z})]^\top \vec{0} && (5.119) \\
&= f(\vec{z}) && (5.120) \\
&= f(\theta \vec{x} + (1-\theta)\vec{y}). && (5.121)
\end{aligned}$$

□

We also have a corresponding second-order condition.

Theorem 124 (Second-Order Condition for Convexity)
Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f : \Omega \to \mathbb{R}$ be a twice-differentiable function. Then f is convex if and only if for all $\vec{x} \in \Omega$, we have

$$\nabla^2 f(\vec{x}) \succeq 0, \tag{5.122}$$

i.e., $\nabla^2 f(\vec{x})$ is PSD for each $\vec{x} \in \Omega$.

Corollary 125. *Let $Q \in \mathbb{S}^n$ be a symmetric matrix, let $\vec{b} \in \mathbb{R}^n$, and let $c \in \mathbb{R}$. The quadratic form*

$$f(\vec{x}) = \vec{x}^\top Q \vec{x} + \vec{b}^\top \vec{x} + c \tag{5.123}$$

is convex if and only if Q is PSD.

Lastly, we identify a strengthened condition of convexity which allows for stronger guarantees later down the line.

Definition 126 (Strictly Convex Function)
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. We say that f is *strictly convex* if the domain of f, say Ω, is convex, and, for any $\vec{x}_1 \neq \vec{x}_2 \in \Omega$ and $\theta \in (0,1)$, we have

$$f(\theta \vec{x}_1 + (1-\theta)\vec{x}_2) < \theta f(\vec{x}_1) + (1-\theta) f(\vec{x}_2). \tag{5.124}$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *strictly concave* if $-f$ is strictly convex.

And correspondingly, we have the first-order and second-order conditions.

Theorem 127 (First-Order Condition for Strict Convexity)
Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f : \Omega \to \mathbb{R}$ be a differentiable function. Then f is strictly convex if and only if for all $\vec{x} \neq \vec{y} \in \Omega$, we have

$$f(\vec{y}) > f(\vec{x}) + [\nabla f(\vec{x})]^\top (\vec{y} - \vec{x}). \tag{5.125}$$

Theorem 128 (Second-Order Condition for Strict Convexity)
Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f : \Omega \to \mathbb{R}$ be a twice-differentiable function such that $\nabla^2 f(\vec{x}) \succ 0$ for all $\vec{x} \in \Omega$, i.e., $\nabla^2 f(\vec{x})$ is PD for all $\vec{x} \in \Omega$. Then f is strictly convex.

Notice that this is *not* an if-and-only-if! As an example, take the scalar function $f(x) = x^4$. Then $f''(0) = 0$, so it is not true that $f''(x) \succ 0$ for all $x \in \Omega = \mathbb{R}$, but f is strictly convex.

5.2.1 Affine Functions

Finally, we spend some time on the special case of *affine* functions, which are the only functions that are both convex and concave. These are of special importance in the remainder of the course, being used or referenced in almost every topic henceforth.

Definition 129 (Affine Functions)

(a) A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *affine* if there exists some vector $\vec{a} \in \mathbb{R}^n$ and some scalar $b \in \mathbb{R}$ such that for any $\vec{x} \in \mathbb{R}^n$, we have
$$f(\vec{x}) = \vec{a}^\top \vec{x} + b. \tag{5.126}$$

(b) A function $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *affine* if there exists some matrix $A \in \mathbb{R}^{m\times n}$ and some vector $\vec{b} \in \mathbb{R}^m$ such that for any $\vec{x} \in \mathbb{R}^n$, we have
$$\vec{f}(\vec{x}) = A\vec{x} + \vec{b}. \tag{5.127}$$

(c) A function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is said to be *affine* if there exists some matrix $A \in \mathbb{R}^{m\times n}$ and scalar $b \in \mathbb{R}$ such that for any $X \in \mathbb{R}^{m\times n}$, we have
$$f(X) = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}X_{ij} + b = \operatorname{tr}\left(A^\top X\right) + b. \tag{5.128}$$

Note that a given function $f : \mathbb{R}^n \to \mathbb{R}$ is affine if and only if the function $g : \mathbb{R}^n \to \mathbb{R}$ defined by $g(\vec{x}) = f(\vec{x}) - f(\vec{0})$ is linear. An analogous result holds for other types of affine functions.

Below, we show that a scalar-valued affine function is one that is both convex and concave, while a vector-valued affine function is one whose component functions are all both convex and concave. Analogous results hold for affine functions whose inputs and outputs are both matrices.

Proposition 130

(a) A function $f : \mathbb{R}^n \to \mathbb{R}$ is affine if and only if it is both convex and concave, i.e., for any $\alpha \in [0, 1]$ and $\vec{x}, \vec{y} \in \mathbb{R}^n$, we have
$$f(\alpha\vec{x} + (1-\alpha)\vec{y}) = \alpha f(\vec{x}) + (1-\alpha) f(\vec{y}). \tag{5.129}$$

(b) A function $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$ is affine if and only each component function of f is both convex and concave, i.e., for any $\alpha \in [0, 1]$ and $\vec{x}, \vec{y} \in \mathbb{R}^n$, we have
$$\vec{f}(\alpha\vec{x} + (1-\alpha)\vec{y}) = \alpha\vec{f}(\vec{x}) + (1-\alpha)\vec{f}(\vec{y}). \tag{5.130}$$

(c) A function $f\colon \mathbb{R}^{m\times n} \to \mathbb{R}$ is affine if and only if it is both convex and concave, i.e., for any $\alpha \in [0,1]$ and $X, Y \in \mathbb{R}^{m\times n}$, we have

$$f(\alpha X + (1-\alpha)Y) = \alpha f(X) + (1-\alpha) f(Y). \tag{5.131}$$

Proof. We prove (a); the claims (b) and (c) follow similarly.

Suppose first that f is affine. Then there exists $\vec{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$ such that for each $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $\alpha \in [0,1]$, we have

$$f(\alpha\vec{x} + (1-\alpha)\vec{y}) = \vec{a}^\top(\alpha\vec{x} + (1-\alpha)\vec{y}) + b \tag{5.132}$$
$$= \alpha\vec{a}^\top\vec{x} + (1-\alpha)\vec{a}^\top\vec{y} + b \tag{5.133}$$
$$= \alpha(\vec{a}^\top\vec{x} + b) + (1-\alpha)(\vec{a}^\top\vec{y} + b) \tag{5.134}$$
$$= \alpha f(\vec{x}) + (1-\alpha) f(\vec{y}). \tag{5.135}$$

Conversely, suppose that

$$f(\alpha\vec{x} + (1-\alpha)\vec{y}) = \alpha f(\vec{x}) + (1-\alpha) f(\vec{y}), \qquad \text{for all } \alpha \in [0,1] \text{ and } \vec{x}, \vec{y} \in \mathbb{R}^n. \tag{5.136}$$

To show that f is affine, it suffices to show that $g\colon \mathbb{R}^n \to \mathbb{R}$ defined by $g(\vec{x}) = f(\vec{x}) - f(\vec{0})$ is linear (and thus can be written as an inner product against a vector $\vec{a}$). We first show that $g(r\vec{x}) = rg(\vec{x})$ for any $r \in \mathbb{R}$ and $\vec{x} \in \mathbb{R}^n$. We break this problem up into three cases, each building on the other.

Case 1. Suppose that $r \in [0,1]$. Then $r\vec{x}$ can be expressed as a convex combination of $\vec{0}$ and $\vec{x}$: that is,

$$r\vec{x} = \alpha\vec{x} + (1-\alpha)\vec{0}, \qquad \text{where } \alpha = r \in (0,1). \tag{5.137}$$

(Yes, this is a simpler step, but we build on it in the later parts.) With this, we have

$$f(r\vec{x}) = f(r\vec{x} + (1-r)\vec{0}) = rf(\vec{x}) + (1-r)f(\vec{0}). \tag{5.138}$$

Thus, we obtain

$$g(r\vec{x}) = f(r\vec{x}) - f(\vec{0}) \tag{5.139}$$
$$= rf(\vec{x}) + (1-r)f(\vec{0}) - f(\vec{0}) \tag{5.140}$$
$$= rf(\vec{x}) - rf(\vec{0}) \tag{5.141}$$
$$= r(f(\vec{x}) - f(\vec{0})) \tag{5.142}$$
$$= rg(\vec{x}). \tag{5.143}$$

Case 2. Now suppose that $r \in (1,\infty)$. Then $\vec{x}$ can be expressed as a convex combination of $\vec{0}$ and $r\vec{x}$: that is,

$$\vec{x} = \alpha(r\vec{x}) + (1-\alpha)\vec{0}, \qquad \text{where } \alpha = \frac{1}{r} \in (0,1). \tag{5.144}$$

Thus we have

$$f(\vec{x}) = f\left(\frac{1}{r}\cdot r\vec{x}\right) = f\left(\frac{1}{r}\cdot r\vec{x} + \left(1 - \frac{1}{r}\right)\vec{0}\right) = \frac{1}{r}f(r\vec{x}) + \left(1-\frac{1}{r}\right)f(\vec{0}). \tag{5.145}$$

Multiplying both sides by r and plugging it into the previous calculation, we get

$$g(r\vec{x}) = f(r\vec{x}) - f(\vec{0}) \tag{5.146}$$

$$= rf(\vec{x}) - r\left(1 - \frac{1}{r}\right) f(\vec{0}) - f(\vec{0}) \tag{5.147}$$

$$= rf(\vec{x}) - (r-1)f(\vec{0}) - f(\vec{0}) \tag{5.148}$$

$$= rf(\vec{x}) - rf(\vec{0}) \tag{5.149}$$

$$= r(f(\vec{x}) - f(\vec{0})) \tag{5.150}$$

$$= rg(\vec{x}). \tag{5.151}$$

Case 3. Now suppose that $r \in (-\infty, 0)$. Then $\vec{0}$ can be expressed as a convex combination of $\vec{x}$ and $r\vec{x}$: that is,

$$0 = \alpha(r\vec{x}) + (1-\alpha)\vec{x}, \text{where } \alpha = \frac{1}{1-r} \in (0,1). \tag{5.152}$$

Thus we have

$$f(\vec{0}) = f\left(\frac{1}{1-r}(r\vec{x}) + \left(1 - \frac{1}{1-r}\right)\vec{x}\right) = \frac{1}{1-r}f(r\vec{x}) + \frac{1-r-1}{1-r}f(\vec{x}) = \frac{1}{1-r}f(r\vec{x}) - \frac{r}{1-r}f(\vec{x}). \tag{5.153}$$

Multiplying both sides by $1-r$ and plugging it into the previous calculation, we get

$$g(r\vec{x}) = f(r\vec{x}) - f(\vec{0}) \tag{5.154}$$

$$= rf(\vec{x}) + (1-r)f(\vec{0}) - f(\vec{0}) \tag{5.155}$$

$$= rf(\vec{x}) - rf(\vec{0}) \tag{5.156}$$

$$= r(f(\vec{x}) - f(\vec{0})) \tag{5.157}$$

$$= rg(\vec{x}). \tag{5.158}$$

This proves that, no matter the value of r, we have $g(r\vec{x}) = rg(\vec{x})$.

Now we show that $g(\vec{x} + \vec{y}) = g(\vec{x}) + g(\vec{y})$. With the preceding result in hand, this proof is much shorter:

$$g(\vec{x} + \vec{y}) = 2g\left(\frac{1}{2}\vec{x} + \frac{1}{2}\vec{y}\right) \tag{5.159}$$

$$= 2\left[f\left(\frac{1}{2}\vec{x} + \frac{1}{2}\vec{y}\right) - f(\vec{0})\right] \tag{5.160}$$

$$= 2\left[\frac{1}{2}f(\vec{x}) + \frac{1}{2}f(\vec{y}) - f(\vec{0})\right] \tag{5.161}$$

$$= 2\left[\frac{1}{2}(f(\vec{x}) - f(\vec{0})) + \frac{1}{2}(f(\vec{y}) - f(\vec{0}))\right] \tag{5.162}$$

$$= 2\left[\frac{1}{2}g(\vec{x}) + \frac{1}{2}g(\vec{y})\right] \tag{5.163}$$

$$= g(\vec{x}) + g(\vec{y}). \tag{5.164}$$

Thus we proved that g is linear, so f is affine. This is a full proof of (a), and (b) and (c) can be proved in almost exactly the same way. □

5.3 Convex Optimization Problems

This section will lay out some of the key properties of the main class of optimization problems we are interested in — convex optimization problems.

Definition 131 (Convex Optimization Problem)
Let $\Omega \subseteq \mathbb{R}^n$ be a set and let $f : \Omega \to \mathbb{R}$ be a function. We say that the problem

$$\min_{\vec{x} \in \Omega} f(\vec{x}) \tag{5.165}$$

is a *convex optimization problem* if Ω is a convex set and f is a convex function.

Note that this applies to other kinds of constraint sets too — in particular, those of the "standard" form "$f_i(\vec{x}) \leq 0$ for all i and $h_i(\vec{x}) = 0$ for all j" still define a feasible region Ω and thus can furnish a convex optimization problem. In particular, we have the following result.

Theorem 132 (When is Standard Form Optimization Problem Convex?)
Let $f_1, \ldots, f_m, h_1, \ldots, h_p : \mathbb{R}^n \to \mathbb{R}$ be functions. The feasible set

$$\Omega = \left\{ \vec{x} \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} f_i(\vec{x}) \leq 0, & \forall i \in \{1, \ldots, m\} \\ h_j(\vec{x}) = 0, & \forall j \in \{1, \ldots, p\} \end{array} \right\} \tag{5.166}$$

is convex if each f_i is convex and each h_j is affine. Consequently, the problem

$$\begin{aligned} \min_{\vec{x} \in \mathbb{R}^n} \quad & f_0(\vec{x}) \\ \text{s.t.} \quad & f_i(\vec{x}) \leq 0, \qquad \forall i \in \{1, \ldots, m\} \\ & h_j(\vec{x}) = 0, \qquad \forall j \in \{1, \ldots, p\}. \end{aligned} \tag{5.167}$$

is a convex optimization problem if $f_0, f_1, \ldots, f_m$ are convex functions and $h_1, \ldots, h_p$ are affine functions.

Proof. Left as exercise. □

Now, we can establish the first-order condition for optimality within a convex problem. This is one of the main theorems of convex analysis.

Theorem 133 (First-Order Conditions for Optimality in Convex Problem)
Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f : \Omega \to \mathbb{R}$ be a differentiable convex function. Let $\vec{x}^\star \in \Omega$ be such that $\nabla f(\vec{x}^\star) = \vec{0}$. Then $\vec{x}^\star \in \operatorname{argmin}_{\vec{x} \in \Omega} f(\vec{x})$, i.e., $\vec{x}^\star$ is a global minimizer of f.

Proof. Let $\vec{y}$ be any other point in Ω. Then

$$\begin{aligned} f(\vec{y}) &\geq f(\vec{x}^\star) + [\nabla f(\vec{x}^\star)]^\top (\vec{y} - \vec{x}^\star) && (5.168) \\ &= f(\vec{x}^\star) + \vec{0}^\top (\vec{y} - \vec{x}^\star) && (5.169) \\ &= f(\vec{x}^\star). && (5.170) \end{aligned}$$

This proves that $\vec{x}^\star \in \operatorname{argmin}_{\vec{x} \in \Omega} f(\vec{x})$ as desired. □

A generalization of this statement is that all local minimizers are global minimizers.

Theorem 134 (For Convex Functions, Local Minima are Global Minima)
Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f : \Omega \to \mathbb{R}$ be a convex function. Let $\vec{x}^\star \in \Omega$ be such that there exists some $\epsilon > 0$ such that if $\vec{x} \in \Omega$ has $\|\vec{x} - \vec{x}^\star\|_2 \leq \epsilon$ then $f(\vec{x}^\star) \leq f(\vec{x})$. Then $\vec{x}^\star$ is a minimizer of f over Ω.

Proof. Left as exercise. □

When is the global minimizer unique? We can justify this using strict convexity.

Theorem 135 (Strictly Convex Functions have Unique Minimizers)
Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f : \Omega \to \mathbb{R}$ be a strictly convex function. Then f has *at most* one global minimizer.

Proof. Left as exercise. □

For an example of a strictly convex function with one global minimizer, take $f(x) = x^4$, which is minimized at $x = 0$. For an example of a strictly convex function with no global minimizers, take $f(x) = e^x$.

5.4 Solving Convex Optimization Problems

To construct a first attempt at systematically solving convex optimization problems, we first need to define active and inactive constraints.

Definition 136 (Types of Constraints)
Consider a problem of the form

$$\begin{aligned} \min_{\vec{x} \in \mathbb{R}^n} \quad & f_0(\vec{x}) \\ \text{s.t.} \quad & f_i(\vec{x}) \leq 0, \qquad \forall i \in \{1, \ldots, m\} \\ & h_j(\vec{x}) = 0, \qquad \forall j \in \{1, \ldots, p\}. \end{aligned} \tag{5.171}$$

Fix a feasible $\vec{x}_0 \in \mathbb{R}^n$. The *inequality* constraint $f_k(\vec{x}) \leq 0$ is *active at* $\vec{x}_0$ if $f_k(\vec{x}_0) = 0$, and *inactive at* $\vec{x}_0$ otherwise, i.e., $f_k(\vec{x}_0) < 0$.

We can use this to formulate a strategy for solving convex optimization problems. Recall that for a convex problem $\min_{\vec{x} \in \Omega} f_0(\vec{x})$ which has a solution $\vec{x}^\star$, either $\nabla f(\vec{x}^\star) = \vec{0}$ or $\vec{x}^\star$ is on the boundary of Ω. The boundary of Ω is any point in which any inequality constraint is active. This allows us to systematically find solutions to convex optimization problems.

Problem Solving Strategy 137. *To solve a convex optimization program, we can do the following.*

1. *Iterate through all 2^m subsets $S \subseteq \{1, \ldots, m\}$ of constraints which might be active at optimum.*
2. *For each S:*
   (i) *Solve the modified problem*

$$\min_{\vec{x} \in \mathbb{R}^n} \quad f_0(\vec{x}) \tag{5.172}$$

$$\begin{aligned} \text{s.t.} \quad & f_i(\vec{x}) = 0, \qquad \forall i \in S \\ & h_j(\vec{x}) = 0, \qquad \forall j \in \{1, \dots, p\}, \end{aligned}$$

i.e., solve the problem where you pretend that all inequality constraints in S are met with equality and pretend that the other inequality constraints don't exist. This gives some solutions $\vec{x}_S^\star$.

(ii) If there is a solution $\vec{x}_S^\star$ which is feasible for the original problem, write down the value of $f_0(\vec{x}_S^\star)$. Otherwise, ignore it.

(iii) After iterating through all $\vec{x}_S^\star$ which are feasible for the original problem, take the one(s) with the best objective value $f_0(\vec{x}_S^\star)$ as the optimal solution(s) to the original problem.

Predictably, this problem solving strategy is exponentially hard as the number of inequality constraints increases. Even if solving the "inner" equality-constrained minimization problems is easy (as it often is), the whole procedure is untenable for large-scale problems. In future chapters, we will develop better analytic and algorithmic ways to solve convex optimization problems.

5.5 Problem Transformations and Reparameterizations

In this section, we discuss various transformations, or reparameterizations, of generic optimization problems, which allow us to go between many equivalent formulations of a problem. Some of them will even allow us to turn non-convex problems into convex problems.

5.5.1 Monotone Transformations of the Objective Function

The core of this technique is the following result.

Proposition 138 (Positive Monotone Transformations Don't Affect Optimizers)
Let Ω be any set, let $f_0\colon \Omega \to \mathbb{R}$ be any function. Define $f_0(\Omega) \doteq \{f_0(\vec{x}) \mid \vec{x} \in \Omega\} \subseteq \mathbb{R}$.

(a) Let $\phi\colon f_0(\Omega) \to \mathbb{R}$ be any monotonically increasing function. Then

$$\operatorname*{argmin}_{\vec{x}\in\Omega} f_0(\vec{x}) = \operatorname*{argmin}_{\vec{x}\in\Omega} \phi(f_0(\vec{x})) \qquad \text{and} \qquad \operatorname*{argmax}_{\vec{x}\in\Omega} f_0(\vec{x}) = \operatorname*{argmax}_{\vec{x}\in\Omega} \phi(f_0(\vec{x})). \tag{5.173}$$

(b) Let $\psi\colon f_0(\Omega) \to \mathbb{R}$ be any monotonically decreasing function. Then

$$\operatorname*{argmin}_{\vec{x}\in\Omega} f_0(\vec{x}) = \operatorname*{argmax}_{\vec{x}\in\Omega} \psi(f_0(\vec{x})) \qquad \text{and} \qquad \operatorname*{argmax}_{\vec{x}\in\Omega} f_0(\vec{x}) = \operatorname*{argmin}_{\vec{x}\in\Omega} \psi(f_0(\vec{x})). \tag{5.174}$$

Proof. We only prove the very first equality; the rest follow similarly. We have

$$\vec{x}^\star \in \operatorname*{argmin}_{\vec{x}\in\Omega} f_0(\vec{x}) \tag{5.175}$$

$$\iff f_0(\vec{x}^\star) \leq f_0(\vec{x}), \qquad \forall \vec{x} \in \Omega \tag{5.176}$$

$$\iff \phi(f_0(\vec{x}^\star)) \leq \phi(f_0(\vec{x})), \qquad \forall \vec{x} \in \Omega \tag{5.177}$$

$$\iff \vec{x}^\star \in \operatorname*{argmin}_{\vec{x}\in\Omega} \phi(f_0(\vec{x})). \tag{5.178}$$

□

The first equality will be by far the most important for us, though the others might also be situationally useful. This proposition is why doing things like squaring the norm in least squares won't affect the solutions we get, i.e.,

$$\operatorname*{argmin}_{\vec{x}\in\mathbb{R}^n} \left\|A\vec{x}-\vec{b}\right\|_2 = \operatorname*{argmin}_{\vec{x}\in\mathbb{R}^n} \left\|A\vec{x}-\vec{b}\right\|_2^2. \tag{5.179}$$

But the fact that ϕ is monotonic *when restricted to* $f_0(\Omega)$ is quite crucial; indeed, we have that

$$\operatorname*{argmin}_{x\in\mathbb{R}} x = \emptyset \qquad \text{but} \qquad \operatorname*{argmin}_{x\in\mathbb{R}} x^2 = \{0\}. \tag{5.180}$$

This is because the function $u \mapsto u^2$ is not monotonic in general, although it is monotonic on the non-negative real numbers $\mathbb{R}_+$. It just so happens that $\|\cdot\|_2$ only outputs non-negative numbers, so actually in the case of least squares we have $f_0(\Omega) = \mathbb{R}_+$ and the proposition applies.

Example 139 (Logistic Regression)**.** A more non-trivial example is logistic regression. First, define $\sigma\colon \mathbb{R} \to (0,1)$ by

$$\sigma(x) \doteq \frac{1}{1+e^{-x}}. \tag{5.181}$$

Suppose we have data points $\vec{x}_1,\dots,\vec{x}_n \in \mathbb{R}^d$ and accompanying labels $y_1,\dots,y_n \in \{0,+1\}$. Suppose that the conditional probability that $y_i = 1$ given $\vec{x}_i$ is given by

$$\mathbb{P}_{\vec{w}_0}[\vec{y}_i = 1 \mid \vec{x}_i] = \sigma(\vec{x}_i^\top \vec{w}_0) \tag{5.182}$$

for some $\vec{w}_0 \in \mathbb{R}^d$. We wish to recover $\vec{w}_0$, and thus recover the generative model $\mathbb{P}_{\vec{w}_0}[y \mid \vec{x}]$. We do this by maximum likelihood estimation. Writing out the problem, we get

$$\begin{aligned}
\vec{w}^\star &\in \operatorname*{argmax}_{\vec{w}\in\mathbb{R}^d} \mathbb{P}_{\vec{w}}[y_1,\dots,y_n \mid \vec{x}_1,\dots,\vec{x}_n] && (5.183)\\
&= \operatorname*{argmax}_{\vec{w}\in\mathbb{R}^d} \prod_{i=1}^n \mathbb{P}_{\vec{w}}[y_i \mid \vec{x}_i] && (5.184)\\
&= \operatorname*{argmax}_{\vec{w}\in\mathbb{R}^d} \left(\prod_{\substack{i=1\\y_i=0}}^n \mathbb{P}_{\vec{w}}[y_i = 0 \,\middle|\, \vec{x}_i]\right)\left(\prod_{\substack{i=1\\y_i=1}}^n \mathbb{P}_{\vec{w}}[y_i = 1 \,\middle|\, \vec{x}_i]\right) && (5.185)\\
&= \operatorname*{argmax}_{\vec{w}\in\mathbb{R}^d} \prod_{i=1}^n \mathbb{P}_{\vec{w}}[y_i = 1 \mid \vec{x}_i]^{y_i}\mathbb{P}_{\vec{w}}[y_i = 0 \mid \vec{x}_i]^{1-y_i} && (5.186)\\
&= \operatorname*{argmax}_{\vec{w}\in\mathbb{R}^d} \prod_{i=1}^n \sigma(\vec{x}_i^\top\vec{w})^{y_i}(1-\sigma(\vec{x}_i^\top\vec{w}))^{1-y_i}. && (5.187)
\end{aligned}$$

Now, the σ function is *very* non-convex. The product of σs is *also* very non-convex. Thus, it seems intractable to solve this problem. But note that the objective function takes values in $(0,1)$. We use the above proposition with the function $x \mapsto \log(x)$, which is monotonically increasing on $(0,1)$. We obtain

$$\begin{aligned}
&\operatorname*{argmax}_{\vec{w}\in\mathbb{R}^d} \prod_{i=1}^n \sigma(\vec{x}_i^\top\vec{w})^{y_i}(1-\sigma(\vec{x}_i^\top\vec{w}))^{1-y_i} && (5.188)\\
&= \operatorname*{argmax}_{\vec{w}\in\mathbb{R}^d} \log\left(\prod_{i=1}^n \sigma(\vec{x}_i^\top\vec{w})^{y_i}(1-\sigma(\vec{x}_i^\top\vec{w}))^{1-y_i}\right) && (5.189)\\
&= \operatorname*{argmax}_{\vec{w}\in\mathbb{R}^d} \sum_{i=1}^n \left(y_i\log\left(\sigma(\vec{x}_i^\top\vec{w})\right) + (1-y_i)\log\left(1-\sigma(\vec{x}_i^\top\vec{w})\right)\right) && (5.190)
\end{aligned}$$

$$= \operatorname*{argmin}_{\vec{w}\in\mathbb{R}^d} \left\{ -\sum_{i=1}^{n} \left(y_i \log\left(\sigma(\vec{x}_i^\top \vec{w})\right) + (1-y_i)\log\left(1-\sigma(\vec{x}_i^\top \vec{w})\right)\right)\right\} \tag{5.191}$$

$$= \operatorname*{argmin}_{\vec{w}\in\mathbb{R}^d} \sum_{i=1}^{n} \left(-y_i \log\left(\sigma(\vec{x}_i^\top \vec{w})\right) - (1-y_i)\log\left(1-\sigma(\vec{x}_i^\top \vec{w})\right)\right). \tag{5.192}$$

In the penultimate line we used another one of the equalities in the proposition with the monotonically decreasing function $\psi(x) = -x$. Thus, logistic regression reduces to minimizing the objective function

$$f_0(\vec{w}) = \sum_{i=1}^{n} \left(-y_i \log\left(\sigma(\vec{x}_i^\top \vec{w})\right) - (1-y_i)\log\left(1-\sigma(\vec{x}_i^\top \vec{w})\right)\right). \tag{5.193}$$

This is the so-called *cross-entropy* loss.

Computing the gradient and Hessian of this function is an exercise, but the result is

$$\nabla f_0(\vec{w}) = \sum_{i=1}^{n} \vec{x}_i(\sigma(\vec{x}_i^\top \vec{w}) - y_i) \tag{5.194}$$

$$\nabla^2 f_0(\vec{w}) = \sum_{i=1}^{n} \sigma(\vec{x}_i^\top \vec{w})(1-\sigma(\vec{x}_i^\top \vec{w}))\vec{x}_i\vec{x}_i^\top. \tag{5.195}$$

Because $\sigma(x) \in (0,1)$, the above Hessian is a non-negative weighted sum of positive semidefinite matrices $\vec{x}_i\vec{x}_i^\top$ and is thus positive semidefinite. By the second order conditions, f_0 is convex! Thus we have turned an extremely non-convex problem into an unconstrained convex minimization problem just by a neat application of monotone functions. We can efficiently solve this problem algorithmically using convex optimization solvers such as gradient descent.

Actually, we can do better than gradient descent for this particular example! If we define

$$X = \begin{bmatrix} \vec{x}_1^\top \\ \vdots \\ \vec{x}_n^\top \end{bmatrix} \in \mathbb{R}^{n\times d}, \qquad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n, \qquad \vec{p}(\vec{w}) = \begin{bmatrix} \sigma(\vec{x}_1^\top \vec{w}) \\ \vdots \\ \sigma(\vec{x}_n^\top \vec{w}) \end{bmatrix} \in \mathbb{R}^n \tag{5.196}$$

then the gradient looks like

$$\nabla f_0(\vec{w}) = X^\top(\vec{p}(\vec{w}) - \vec{y}). \tag{5.197}$$

Notice the similarity to the gradient of least squares, $X^\top(X\vec{w} - \vec{y})$. In fact, we can exploit this similarity to obtain a specialized optimization algorithm called *iteratively reweighted least squares*, which can efficiently solve for maximum likelihood estimates within this type of statistical model (i.e., a *generalized linear model*).

5.5.2 Slack Variables

The purpose of slack variables is to turn equality constraints into inequality constraints and vice-versa.

Definition 140 (Slack Variables)
Consider a problem of the form

$$\begin{aligned} \min_{\vec{x}\in\mathbb{R}^n} \quad & f_0(\vec{x}) \\ \text{s.t.} \quad & f_i(\vec{x}) \leq 0, \qquad \forall i \in \{1,\dots,m\} \\ & h_j(\vec{x}) = 0, \qquad \forall j \in \{1,\dots,p\}. \end{aligned} \tag{5.198}$$

Let $\mathcal{S} \subseteq \{1, \ldots, m\}$ be any subset. The above problem is equivalent to the following problem:

$$\begin{aligned} \min_{\substack{\vec{x} \in \mathbb{R}^n \\ \vec{s} \in \mathbb{R}^{\mathcal{S}}_+}} \quad & f_0(\vec{x}) \\ \text{s.t.} \quad & f_i(\vec{x}) + s_i = 0, \qquad \forall i \in \mathcal{S} \\ & f_i(\vec{x}) \leq 0, \qquad \forall i \in \{1, \ldots, m\} \setminus \mathcal{S} \\ & h_j(\vec{x}) = 0, \qquad \forall j \in \{1, \ldots, p\}. \end{aligned} \tag{5.199}$$

Here the notation $\mathbb{R}^{\mathcal{S}}_+ = \{(x_i)_{i \in \mathcal{S}} \mid x_i \geq 0 \; \forall i \in \mathcal{S}\}$, and $\vec{s}$ is called a *slack variable*.

One can choose to create slack variables s_i for only a subset of the inequality constraints, or all of them. When we work with more advanced optimization algorithms later, sometimes this parameterization is crucial (e.g. for equality-constrained Newton's method).

Example 141. If we have a problem of the form

$$\min_{\vec{x} \in \mathbb{R}^n} \quad f_0(\vec{x}) \tag{5.200}$$

$$\text{s.t.} \quad 3x_1^2 + 4x_2^2 \leq 0 \tag{5.201}$$

$$2x_1^2 + 5x_2^2 \leq 0 \tag{5.202}$$

but our solver could only handle equality constraints, then it would be equivalent to solve the problem

$$\min_{\substack{\vec{x} \in \mathbb{R}^2 \\ \vec{s} \in \mathbb{R}^2_+}} \quad f_0(\vec{x}) \tag{5.203}$$

$$\text{s.t.} \quad 3x_1^2 + 4x_2^2 + s_1 = 0 \tag{5.204}$$

$$2x_1^2 + 5x_2^2 + s_2 = 0 \tag{5.205}$$

and, upon solving this problem and obtaining $(\vec{x}^\star, \vec{s}^\star)$, the solution to the original problem would be this same $\vec{x}^\star$.

5.5.3 Epigraph Reformulation

The epigraph reformulation is a way to convert between non-linearity in the objective and a constraint.

Definition 142

Consider a problem of the form

$$\begin{aligned} \min_{\vec{x} \in \mathbb{R}^n} \quad & f_0(\vec{x}) \\ \text{s.t.} \quad & f_i(\vec{x}) \leq 0, \qquad \forall i \in \{1, \ldots, m\} \\ & h_j(\vec{x}) = 0, \qquad \forall j \in \{1, \ldots, p\}. \end{aligned} \tag{5.206}$$

Its *epigraph reformulation* is the problem

$$\begin{aligned} \min_{\substack{t \in \mathbb{R} \\ \vec{x} \in \mathbb{R}^n}} \quad & t \\ \text{s.t.} \quad & t \geq f_0(\vec{x}) \\ & f_i(\vec{x}) \leq 0, \qquad \forall i \in \{1, \ldots, m\} \end{aligned} \tag{5.207}$$

$$h_j(\vec{x}) = 0, \qquad \forall j \in \{1, \ldots, p\}.$$

The epigraph objective is always a linear and differentiable function of the decision variables $(t, \vec{x})$. However, the constraint can become complicated if $f_0(\vec{x})$ is non-linear. This transformation is especially useful in the case of quadratically-constrained quadratic programs (QCQP).

Example 143 (Elastic-Net Regularization)**.** This example uses the two previously-discussed techniques in tandem to figure out how to handle a regularizer with both smooth and non-smooth components (i.e., ℓ^1 and ℓ^2 norms).

Let $A \in \mathbb{R}^{m \times n}$, and $\vec{y} \in \mathbb{R}^n$. Suppose that we have a problem of the form

$$\min_{\vec{x} \in \mathbb{R}^n} \left\{ \|A\vec{x} - \vec{y}\|_2^2 + \alpha \|\vec{x}\|_2^2 + \beta \|\vec{x}\|_1 \right\}. \tag{5.208}$$

The regularizer $\alpha \|\vec{x}\|_2^2 + \beta \|\vec{x}\|_1$ is called the *elastic net regularizer* and encourages "sparse" and small $\vec{x}$; this regularizer has some use in the analysis of high-dimensional and structured data.

Suppose that our solver can only handle differentiable objectives, but is able to handle constraints so long as they are also differentiable. Then we cannot solve the problem out-right using our solver, so we need to reformulate it. We can first start by using a modification of the epigraph reformulation:

$$\min_{\substack{t \in \mathbb{R} \\ \vec{x} \in \mathbb{R}^n}} \quad \|A\vec{x} - \vec{y}\|_2^2 + \alpha \|\vec{x}\|_2^2 + \beta t \tag{5.209}$$

$$\text{s.t.} \quad t \geq \|\vec{x}\|_1 . \tag{5.210}$$

Now the constraint is non-differentiable, so we are no longer able to exactly solve this constrained problem. However, the objective is now convex and differentiable. Let us rewrite this problem using the $|x_i|$:

$$\min_{\substack{t \in \mathbb{R} \\ \vec{x} \in \mathbb{R}^n}} \quad \|A\vec{x} - \vec{y}\|_2^2 + \alpha \|\vec{x}\|_2^2 + \beta t \tag{5.211}$$

$$\text{s.t.} \quad t \geq \sum_{i=1}^{n} |x_i| . \tag{5.212}$$

The main insight that goes into resolving the non-differentiability of this constraint is that $|x_i| = \max\{x_i, -x_i\}$ and in particular $s_i \geq |x_i|$ if and only if $s_i \geq x_i$ and $s_i \geq -x_i$. Thus we add more "slack-type" variables and obtain

$$\min_{\substack{t \in \mathbb{R} \\ \vec{x} \in \mathbb{R}^n \\ \vec{s} \in \mathbb{R}^n_+}} \quad \|A\vec{x} - \vec{y}\|_2^2 + \alpha \|\vec{x}\|_2^2 + \beta t \tag{5.213}$$

$$\text{s.t.} \quad t \geq \sum_{i=1}^{n} s_i \tag{5.214}$$

$$s_i \geq x_i, \qquad \forall i \in \{1, \ldots, n\} \tag{5.215}$$

$$s_i \geq -x_i, \qquad \forall i \in \{1, \ldots, n\}. \tag{5.216}$$

Now the constraints are all differentiable (in fact, affine) and the problem may be solved. This problem is exactly equivalent to the above problem because t is being minimized in the objective, so each s_i is being minimized by way of the first constraint, meaning that $s_i = |x_i|$ and this is equivalent to the original elastic-net regression problem.

Chapter 6

Gradient Descent

Relevant sections of the textbooks:

- [2] Section 12.2.

6.1 Strong Convexity and Smoothness

In this section, we will introduce two properties of functions that will become useful in analyzing optimization algorithms. The first property is a notion of convexity of functions that is stronger than the ones previously introduced.

Definition 144 (μ-Strongly Convex Function)
Let $\mu \geq 0$. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. We say that f is *μ-strongly convex* if the domain of f, say Ω, is convex, and, for any $\vec{x}, \vec{y} \in \Omega$, we have

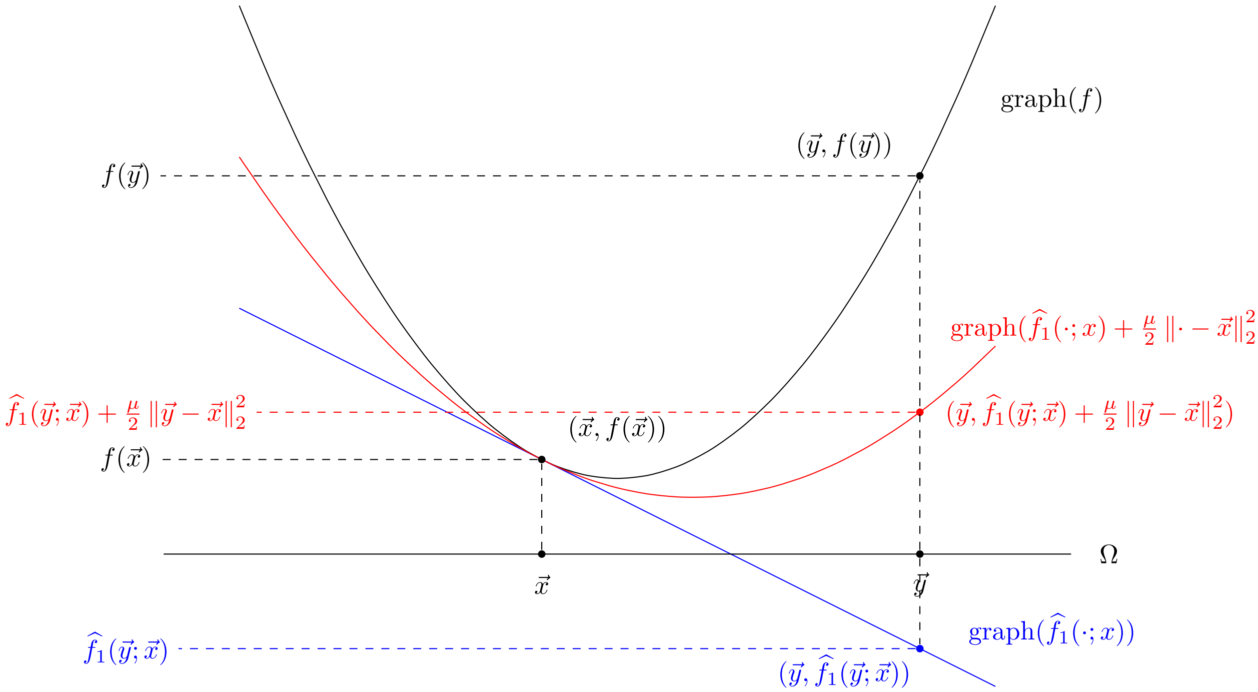
$$f(\vec{y}) \geq f(\vec{x}) + [\nabla f(\vec{x})]^\top (\vec{y} - \vec{x}) + \frac{\mu}{2} \|\vec{y} - \vec{x}\|_2^2 . \tag{6.1}$$

A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is *μ-strongly concave* if $-f$ is μ-strongly convex.

Recall from Theorem 123 that the first order condition for (usual) convexity requires the function to be bigger than its linear (first order) Taylor approximation centered at any point. μ-strong convexity imposes a stronger requirement on the function: it needs to be bigger than its linear approximation plus a non-negative quadratic term that has a Hessian matrix μI. This becomes more obvious if we write Equation (6.1) in the equivalent form

$$f(\vec{y}) \geq \underbrace{f(\vec{x}) + [\nabla f(\vec{x})]^\top (\vec{y} - \vec{x})}_{\text{first-order Taylor approximation}} + \underbrace{(\vec{y} - \vec{x})^\top \left(\frac{\mu}{2} I\right) (\vec{y} - \vec{x})}_{\text{non-negative quadratic term}} . \tag{6.2}$$

Below, we visualize the μ-strong convexity property and compare it to the first order condition for convexity.



Therefore, μ-strong convexity of the function is a very important feature. It guarantees that the function will always have enough *curvature* (at least as much as its quadratic lower bound) and thus will never become too flat anywhere on its domain.

If the function f is twice-differentiable we can formalize this notion by giving the following equivalent condition for μ-strong convexity.

Theorem 145 (Second Order Condition for μ-Strong Convexity)
Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f\colon \Omega \to \mathbb{R}$ be a twice-differentiable function. Then f is μ-strongly convex if and only if for all $\vec{x} \in \Omega$ we have

$$\nabla^2 f(\vec{x}) - \mu I \succeq 0, \tag{6.3}$$

i.e., $\nabla^2 f(\vec{x}) - \mu I$ is PSD for each $\vec{x} \in \Omega$.

An important property of μ-strongly convex functions is that they are strictly convex and thus they have at most one minimizer. In fact, one can show that they have exactly one minimizer.

Theorem 146 (Strongly Convex Functions have Unique Global Minimizers)
Let $\mu > 0$. Let $\Omega \subseteq \mathbb{R}^n$ be a convex set and let $f\colon \Omega \to \mathbb{R}$ be a μ-strongly convex function. Then f has *exactly one* global minimizer.

The second property we want to introduce is L-smoothness, which describes quadratic *upper* bounds on the function f.

Definition 147 (L-Smooth Function)
Let $L \geq 0$. Let $\Omega \subseteq \mathbb{R}^n$ be a set. Let $f\colon \Omega \to \mathbb{R}$ be a differentiable function. We say that f is *L-smooth* if, for

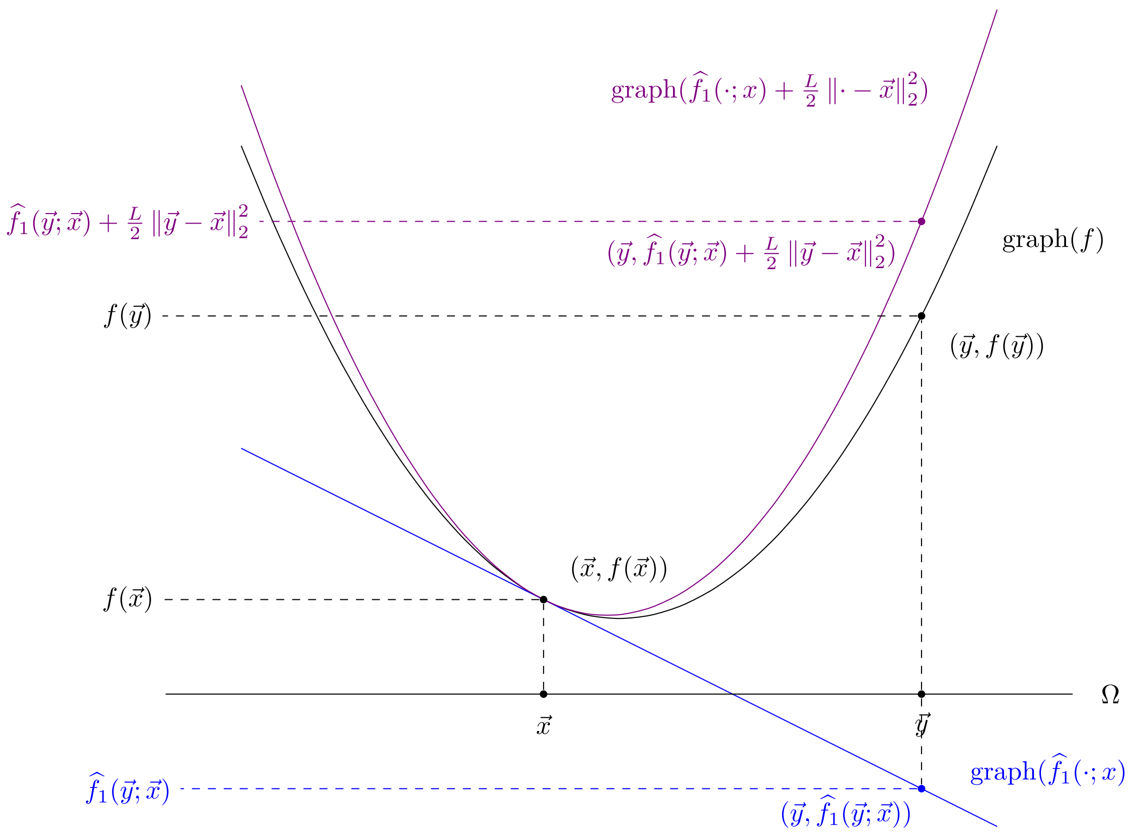
any $\vec{x}, \vec{y} \in \Omega$, we have

$$f(\vec{y}) \leq f(\vec{x}) + [\nabla f(\vec{x})]^\top (\vec{y} - \vec{x}) + \frac{L}{2} \|\vec{y} - \vec{x}\|_2^2 . \tag{6.4}$$

If f is L-smooth, then the function f is *upper* bounded by its first order Taylor approximation plus a non-negative quadratic term:

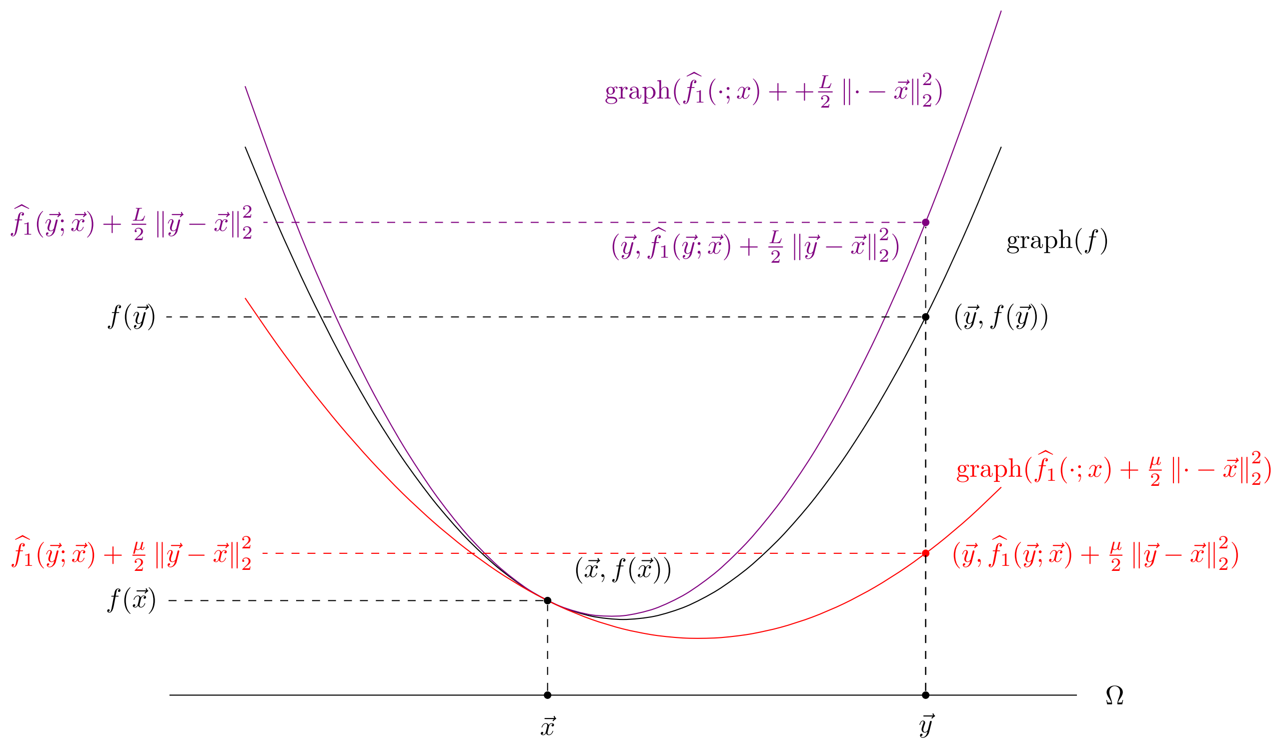
$$f(\vec{y}) \leq \underbrace{f(\vec{x}) + [\nabla f(\vec{x})]^\top (\vec{y} - \vec{x})}_{\text{first-order Taylor approximation}} + \underbrace{(\vec{y} - \vec{x})^\top \left(\frac{L}{2} I\right) (\vec{y} - \vec{x})}_{\text{non-negative quadratic term}} . \tag{6.5}$$

We can visualize the L-smoothness condition similarly to the strongly convex condition, as below:



L-smoothness provides a quadratic upper bound on f. This upper bound ensures that the function doesn't have too much curvature anywhere on its domain (at most as much as its upper bound). We will see later in this chapter that this actually translates into an upper bound on the rate at which the gradient of the function changes.

Finally, we visualize the behavior of the μ-strongly convex and L-smooth bounds together.



6.2 Gradient Descent

We now have all the tools we need to introduce, understand, and analyze gradient descent - one of the most ubiquitous optimization algorithms. The algorithm is used to numerically solve differentiable and unconstrained optimization problems.

More formally, let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. We attempt to solve the problem:

$$p^\star = \min_{\vec{x}\in\mathbb{R}^n} f(\vec{x}). \tag{6.6}$$

The general idea behind the gradient descent algorithm is that it starts with some *initial guess* $\vec{x}_0 \in \mathbb{R}^n$ and produces a sequence of refined guesses $\vec{x}_1, \vec{x}_2, \ldots$, called *iterates*. In each iteration $t = 0, 1, 2, \ldots,$ the algorithm updates its guess according to the following rule:

$$\vec{x}_{t+1} = \vec{x}_t + \eta\vec{v}_t \tag{6.7}$$

for some $\vec{v}_t \in \mathbb{R}^n$ and $\eta \geq 0$.

There are two quantities that we need to specify to complete the definition of the algorithm:

- The vector $\vec{v}_t$, or the *search direction*, which specifies a good direction to move.
- The scalar η, or the *step size*, which specifies how far we move in the direction of $\vec{v}_t$.

For the gradient descent algorithm, we assume that at every point $\vec{x} \in \mathbb{R}^n$ we can get two pieces of information about the function we are optimizing: the value of the function $f(\vec{x}) \in \mathbb{R}$ as well as its gradient $\nabla f(\vec{x}) \in \mathbb{R}^n$. Next, we will use this available information to come up with a good search direction $\vec{v}_t$. The choice of the step size η is a more difficult task. There is no universal choice of η that is good for all problems and a good choice of η is problem-specific. We will discuss the choice of the step size later in the section and show the important role it plays in the algorithm.