Prerequisites

Calculus (multivariate functions)

Linear algebra (norms)

Sparse regression via the lasso
Convex functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is **convex** if for any $x, y \in \mathbb{R}^n$ and any $\theta \in (0, 1)$

$$\theta f(x) + (1 - \theta) f(y) \geq f(\theta x + (1 - \theta) y)$$
Convex functions

\[ \theta f(x) + (1 - \theta)f(y) \]

\[ f(\theta x + (1 - \theta)y) \]
Strictly convex functions

A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is strictly convex if for any $x, y \in \mathbb{R}^n$ and any $\theta \in (0, 1)$

$$\theta f(x) + (1 - \theta) f(y) > f(\theta x + (1 - \theta) y)$$
Strictly convex functions

\[ f(\theta x + (1 - \theta)y) = \theta f(x) + (1 - \theta) f(y) \]
Linear functions

Linear functions are convex

\[ f (\theta x + (1 - \theta) y) = \theta f(x) + (1 - \theta) f(y) \]
Quadratic forms

Let $A$ be a symmetric matrix, if

$$f(x) := x^T Ax \geq 0 \quad \text{for all } x$$

then the quadratic form $f$ is positive semidefinite.

If

$$f(x) := x^T Ax > 0 \quad \text{for all } x$$

then the quadratic form $f$ is positive definite.
Positive semidefinite quadratic forms are convex

\[ \theta f(x) + (1 - \theta)f(y) - f(\theta x + (1 - \theta) y) \]
\[ = \theta x^T Ax + (1 - \theta)y^T Ay - (\theta x + (1 - \theta) y)^T A(\theta x + (1 - \theta) y) \]
\[ = (\theta - \theta^2)x^T Ax + (1 - \theta - (1 - \theta)^2)y^T Ay - 2\theta(1 - \theta)x^T Ay \]
\[ = \theta(1 - \theta)x^T Ax + \theta(1 - \theta)y^T Ay - 2\theta(1 - \theta)x^T Ay \]
\[ = \theta(1 - \theta)(x - y)^T A(x - y) \]

Function is convex if quadratic form is positive semidefinite, strictly convex if it is positive definite
Positive semidefinite quadratic function
Norms are convex

For any \( x, y \in \mathbb{R}^n \) and any \( \theta \in (0, 1) \)

\[
\|\theta x + (1 - \theta) y\| \leq \|\theta x\| + \|(1 - \theta) y\|
\]

\[
= \theta \|x\| + (1 - \theta) \|y\|
\]
\( \ell_0 \) "norm" is not convex

Let \( x := (1, 0) \) and \( y := (0, 1) \), for any \( \theta \in (0, 1) \)

\[
\| \theta x + (1 - \theta) y \|_0 = 2
\]

\[
\theta \| x \|_0 + (1 - \theta) \| y \|_0 = 1
\]
Is the lasso cost function convex?

$f$ strictly convex, $g$ convex, $h := f + \lambda g$?

\[
    h(\theta x + (1 - \theta) y) = f(\theta x + (1 - \theta) y) + \lambda g(\theta x + (1 - \theta) y) \\
    < \theta f(x) + (1 - \theta) f(y) + \lambda \theta g(x) + \lambda (1 - \theta) g(y) \\
    = \theta h(x) + (1 - \theta) h(y)
\]
Lasso cost function is convex

**Sum** of convex functions is convex

If at least one is strictly convex, then sum is strictly convex

**Scaling** by a positive factor preserves convexity

Lasso cost function is convex!
Local minima are global

Any local minimum of a convex function is also a global minimum
Proof

Let $x_{\text{loc}}$ be a local minimum: for all $x \in \mathbb{R}^n$ such that $||x - x_{\text{loc}}||_2 \leq \gamma$

$$f(x_{\text{loc}}) \leq f(x)$$

Let $x_{\text{glob}}$ be a global minimum

$$f(x_{\text{glob}}) < f(x_{\text{loc}})$$
Proof

Choose \( \theta \) so that \( x_\theta := \theta x_{loc} + (1 - \theta) x_{glob} \) satisfies

\[
||x_\theta - x_{loc}||_2 \leq \gamma
\]

then

\[
f(x_{loc}) \leq f(x_\theta) \\
= f(\theta x_{loc} + (1 - \theta) x_{glob}) \\
\leq \theta f(x_{loc}) + (1 - \theta) f(x_{glob}) \quad \text{by convexity of } f \\
< f(x_{loc}) \quad \text{because } f(x_{glob}) < f(x_{loc})
\]
Strictly convex functions

Strictly convex functions have at most one global minimum

Proof: Assume two minima exist at \( x \neq y \) with value \( v_{min} \)

\[
f (0.5x + 0.5y) < 0.5f(x) + 0.5f(y) = v_{min}
\]
What have we learned?

Definition of convexity

The lasso function is convex

The local minima of convex functions are global minima (and are unique for strictly convex functions)