

1) Beck Exercise 7.26. Let C be a strongly convex subset of \mathbb{R}^n . A function f is called strongly convex over C if there exists $\sigma > 0$ such that the function $f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C . The parameter σ is called the strong convexity parameter. In the following questions C is a given subset of \mathbb{R}^n .

i) Prove that f is strongly convex with parameter σ if and only if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{\sigma}{2}\lambda(1 - \lambda)\|x - y\|^2$$

for any $x, y \in C$ and $\lambda \in [0, 1]$.

Sufficient Condition:

Assume the inequality holds. We need to show that f is strongly convex with parameter σ . We will use the definition of strong convexity.

By the given inequality, for any $x, y \in C$ and $\lambda \in [0, 1]$, we have:

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) - \frac{\sigma}{2}\lambda(1 - \lambda)\|x - y\|^2$$

Now, we will relate this to the strong convexity definition of f . A function f is strong convex over C if there exists $\sigma > 0$ such that $f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C . The convexity of this modified function implies that for any $x, y \in C$ and $\lambda \in [0, 1]$, the following holds:

$$f(\lambda x + (1 - \lambda)y) - \frac{\sigma}{2}\|\lambda x + (1 - \lambda)y\|^2 \leq \lambda(f(x) - \frac{\sigma}{2}\|x\|^2) + (1 - \lambda)(f(y) - \frac{\sigma}{2}\|y\|^2)$$

This inequality can be manipulated to resemble the given inequality, so f is strongly convex.

Necessary Condition:

Assume f is strongly convex with parameter σ . We need to show the inequality holds. Given the strong convexity of f , for any $x, y \in C$ and $\lambda \in [0, 1]$, the convexity condition of the modified function $f(x) - \frac{\sigma}{2}\|x\|^2$ holds. From this, the desired inequality can be derived.

ii) Prove that a strongly convex function over C is also strictly convex over C .

Assume f is strongly convex over C with a strong convexity parameter $\sigma > 0$. We need to show that f is also strictly convex over C .

Given the strong convexity of f , for any $x, y \in C$ with $x \neq y$ and $\lambda \in (0, 1)$, the function $g(x) = f(x) - \frac{\sigma}{2}\|x\|^2$ is convex. Therefore, the following inequality holds due to the convexity of g :

$$g(\lambda x + (1 - \lambda)y) \leq \lambda g(x) + (1 - \lambda)g(y)$$

Expanding this inequality gives:

$$f(\lambda x + (1 - \lambda)y) - \frac{\sigma}{2}\|\lambda x + (1 - \lambda)y\|^2 \leq \lambda(f(x) - \frac{\sigma}{2}\|x\|^2) + (1 - \lambda)(f(y) - \frac{\sigma}{2}\|y\|^2)$$

We can expand and simplify this further. Since $\sigma > 0$ and $x \neq y$, the term $\|\lambda x + (1 - \lambda)y\|^2$ is strictly positive. This means that this inequality:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

must hold for any $\lambda \in (0, 1)$ and $x, y \in C$ with $x \neq y$, since the strict inequality comes from the strictly positive term involving σ and the non-linearity of $\|x\|^3$.

Thus f is strictly convex over C .

iii) Suppose that f is continuously differentiable over C . Prove that f is strongly convex over C with parameter σ if and only if

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{\sigma}{2}\|x - y\|^2$$

for any $x, y \in C$.

Sufficient Condition:

Assume the inequality holds for all $x, y \in C$. We want to show that f is strongly convex with parameter σ . Recall that a function is strongly convex over C if the function $g(x) = f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C .

The given inequality can be written as:

$$f(y) - f(x) \geq \nabla f(x)^T(y - x) + \frac{\sigma}{2}\|x - y\|^2$$

This inequality resembles the first-order condition for the convexity of f , but with an additional term involving σ that strengthens the inequality. This additional term implies that the function $g(x)$ is convex, thereby showing that $f(x)$ is strongly convex with parameter σ .

Necessary Condition:

Assume f is strongly convex over C with parameter σ . We need to prove the inequality holds for all $x, y \in C$. From the strong convexity of f , we know that the function $g(x) = f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C .

The convexity of $g(x)$ implies that for any $x, y \in C$, the following inequality holds (using the first-order condition for convexity):

$$g(y) \geq g(x) + \nabla g(x)^T(y - x)$$

Replacing $g(x)$ and $g(y)$ with their definitions gives:

$$f(y) - \frac{\sigma}{2}\|y\|^2 \geq f(x) - \frac{\sigma}{2}\|x\|^2 + (\nabla f(x) - \sigma x)^T(y - x)$$

Rearranging this gives:

$$f(y) \geq f(x) + \nabla f(x)^T(y - x) + \frac{\sigma}{2}\|x - y\|^2$$

Giving us our desired inequality.

iv) Suppose that f is continuously differentiable over C . Prove that f is strongly convex over C with parameter σ if and only if

$$(\nabla f(x) - \nabla f(y))^T(x - y) \geq \sigma\|x - y\|^2$$

for any $x, y \in C$.

Sufficient Condition:

Assume the inequality holds for all $x, y \in C$. We need to show that f is strongly convex with parameter σ . Recall that a function is strongly convex over C if the function $g(x) = f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C .

Given the inequality, for any $x, y \in C$, we have:

$$(\nabla f(x) - \nabla f(y))^T(x - y) \geq \sigma\|x - y\|^2$$

This implies that the gradient of f increases at least linearly with the distance between x and y . This property is a stronger version of the monotonicity of the gradient, which is a characteristic of convex functions. Therefore, this implies that f is strongly convex with the parameter σ providing a lower bound on the curvature of f .

Necessary Condition:

Assume f is strongly convex over C with parameter σ . We need to prove the inequality holds for all $x, y \in C$. From the strong convexity of f , we know that the function $g(x) = f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C .

The convexity of $g(x)$ implies that for any $x, y \in C$, the following inequality holds (using the first-order condition for convexity):

$$g(y) \geq g(x) + \nabla g(x)^T(y - x)$$

substituting $g(x)$ and $g(y)$ with their definitions, we can get the required inequality. Specifically, the strong convexity of f implies a certain lower bound on the difference of gradients, which is given by the inequality:

$$(\nabla f(x) - \nabla f(y))^T(x - y) \geq \sigma\|x - y\|^2$$

v) Suppose that f is twice continuously differentiable over C . Prove that f is strongly convex over C with parameter σ if and only if $\nabla^2 f(x) \succeq \sigma I$ for any $x \in C$.

Sufficient Condition:

Assume that for any $x \in C$, the Hessian of f , $\nabla^2 f(x)$, satisfies $\nabla^2 f(x) \succeq \sigma I$. This means that the eigenvalues of $\nabla^2 f(x)$ are greater than or equal to σ . To show that f is strongly convex with parameter σ , we need to demonstrate that $f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C .

The convexity of a twice continuously differentiable function can be determined by its Hessian. Specifically, a function f is convex over C if $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$. In this case, since $\nabla^2 f(x) \succeq \sigma I$, it implies that $\nabla^2 f(x)$ is indeed positive semidefinite for all $x \in C$, and thus f is convex over C . Furthermore, the condition $\nabla^2 f(x) \succeq \sigma I$ implies a certain level of positive curvature, which implies strong convexity.

Necessary Condition:

Assume f is strongly convex over C with parameter σ . We need to show $\nabla^2 f(x) \succeq \sigma I$ for all $x \in C$. By the definition of strong convexity, $f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C .

Since f is twice continuously differentiable, we can use the Hessian to determine convexity. For f to be strongly convex, its curvature must exceed a

certain threshold, reflected by the parameter σ . This implies that the Hessian of f , $\nabla^2 f(x)$, must not just be positive semidefinite, but also have eigenvalues that are at least σ , which is the same as $\nabla^2 f(x) \succeq \sigma I$.

2) Let $c \in \mathbb{R}^n \setminus \{0\}$ and $A \succ 0$. Consider the problem of minimizing a linear function over the ellipsoid,

$$\min_{x \in \mathbb{R}^n} c^T x$$

such that $x^T A x \leq 1$.

a) Show that this optimization problem is convex.

To show that the given optimization problem is convex, we need to establish that both the objective function and the constraint set are convex.

Convexity of the objective function:

The objective function is $c^T x$, which is a linear function in x . Linear functions are both convex and concave. This is because for any two points $x_1, x_2 \in \mathbb{R}^n$ and any $\lambda \in [0, 1]$, a linear function satisfies:

$$c^T(\lambda x_1 + (1 - \lambda)x_2) = \lambda c^T x_1 + (1 - \lambda)c^T x_2$$

This equality directly satisfies the definition of convexity for a function. Therefore, the objective function $c^T x$ is convex.

Convexity of the constraint set:

The constraint set is defined by $x^T A x \leq 1$. This is a quadratic form, and we need to determine if this set forms a convex set.

Given that $A \succ 0$, the quadratic form $x^T A x$ represents an ellipsoid. For positive definite matrices, the associated quadratic forms are convex. This can be shown using the definition of convexity:

For any x_1, x_2 in the set and any $\lambda \in [0, 1]$,

$$(\lambda x_1 + (1 - \lambda)x_2)^T A (\lambda x_1 + (1 - \lambda)x_2) \leq 1$$

This inequality holds because the left-hand side is a convex combination of points inside the ellipsoid defined by $x^T A x \leq 1$, due to the positive definiteness of A . Thus, the constraint set is convex.

b) Write the solution. Hint: use the change of variables $y = A^{\frac{1}{2}} x$.

Change of variables:

Using the change of variables $y = A^{\frac{1}{2}} x$, we can rewrite the constraint and objective function:

Constraint: The constraint $x^T A x \leq 1$ becomes $(A^{\frac{1}{2}} x)^T (A^{\frac{1}{2}} x) \leq 1$, which is the same as: $y^T y \leq 1$. This describes a unit ball in \mathbb{R}^n .

Objective function: The objective function $c^T x$ can be rewritten in terms of y as $c^T (A^{-\frac{1}{2}} y)$. This is because $x = A^{-\frac{1}{2}} y$.

The optimization problem in y coordinates becomes:

$$\min_{y \in \mathbb{R}^n} c^T A^{-\frac{1}{2}} y$$

with constraint

$$y^T y \leq 1$$

Solving the problem:

The problem now is to minimize a linear function over the unit ball defined by $y^T y \leq 1$. The solution to this problem lies on the boundary of the unit ball since the objective function is linear and we are minimizing over a convex set.

To find the minimum of $c^T A^{-\frac{1}{2}}$ over the unit ball, we need to minimize the dot product of $c^T A^{-\frac{1}{2}}$ with y subject to the constraint $\|y\| = 1$. The minimum value occurs when y is in the direction of $-A^{-\frac{1}{2}}c$. Therefore, the optimal y is:

$$y^* = -\frac{A^{-\frac{1}{2}}c}{\|A^{-\frac{1}{2}}c\|}$$

which is the unit vector in the direction of $-A^{-\frac{1}{2}}c$.

Converting back to the original variables:

To find the optimal x , we convert back using $x = A^{-\frac{1}{2}}y$. Thus:

$$x^* = A^{-\frac{1}{2}}y^* = -\frac{A^{-1}c}{\|A^{-\frac{1}{2}}c\|}$$

Thus, the solution to the optimization problem is:

$$x^* = -\frac{A^{-1}c}{\|A^{-\frac{1}{2}}c\|}.$$

3) Beck Exercise 8.2. Let $C = B[x_0, r]$, where $x_0 \in \mathbb{R}^n$ and $r > 0$ are given. Find a formula for the orthogonal projection operator P_C .

To find the formula for the orthogonal projection operator P_C onto a ball $C = B[x_0, r]$ in \mathbb{R}^n , where $B[x_0, r]$ is the ball centered at x_0 with radius $r > 0$, we need to consider two cases for a given point $y \in \mathbb{R}^n$.

If y is inside the ball or on the boundary:

The orthogonal projection of y onto the ball is y itself since y is already the closest point to itself in the ball.

If y is outside the ball:

The orthogonal projection of y onto the ball is the point on the boundary of the ball that lies on the line segment connecting y and the center x_0 of the ball.

The projection onto the boundary of the ball can be found by moving from x_0 towards y a distance of r . This can be done by normalizing the vector $y - x_0$ to have length r and then translating it so its origin is at x_0 .

Thus, the formula for the orthogonal projection operator is:

$$\begin{cases} x_0 + r \frac{y - x_0}{\|y - x_0\|} & \text{if } \|y - x_0\| > r \\ y & \text{otherwise} \end{cases}$$

4)

The optimal w is: [22.15469572, -10.64394485].
 The optimal b is: -5.289810516097577.
 The hyperplane can be gotten using $w^T x + b = 0$.
 The code to get the optimal w and b is:

```
import cvxpy as cp
import numpy as np

# Generate the random points like the given Matlab code
np.random.seed(314)
x = np.random.rand(40, 1)
y = np.random.rand(40, 1)
classes = (2 * x < y + 0.5) + 1
A1 = np.hstack([x[classes.flatten() == 1], y[classes.flatten() == 1]])
A2 = np.hstack([x[classes.flatten() == 2], y[classes.flatten() == 2]])

# Define the optimization variables
w = cp.Variable(2)
b = cp.Variable()

# Define the constraints
constraints = [w.T @ A1[i] + b >= 1 for i in range(A1.shape[0])]
constraints += [w.T @ A2[i] + b <= -1 for i in range(A2.shape[0])]

# Define the objective function
objective = cp.Minimize(0.5 * cp.norm(w, 2)**2)

# Define the problem and solve
problem = cp.Problem(objective, constraints)
problem.solve()

# Extract the solution
w_value = w.value
b_value = b.value

print("The optimal w is:", w_value)
print("The optimal b is:", b_value)
```