Q: Let $f(\cdot)$ be a convex function defined on domain D. Do you require D to satisfy some property?

$$f(\lambda x_1 + (1-\lambda) x_2) < \lambda_1 f(x_1) + (1-\lambda) f(x_2)$$
$\forall \lambda \in [0,1]$

Ans: 1 D should be continuous

H/W (Proof by contradiction)

2 D should be convex

A set is S is convex iff $\forall x_1, x_2 \in S$

$$\lambda x_1 + (1-\lambda) x_2 \in S \quad \forall \lambda \in [0,1]$$

Back to optimization.

$$\min \quad f(x)$$
$$\text{s.t} \quad g_1(x) \le 0$$

To solve this problem:

1) You might look at the shapes (nature) of these functions (eg. convexity)
2) Characterise subset of search space that is guaranteed to contain point(s) of optimality (if possible such as in linear programs)
3) If domain is open & f is cts you could

consider looking at first & second order necessary (or sufficient) conditions for optimality. However if domain is closed, the gradient & Hessian are not defined on boundary. [Domain is defined by $\{x \mid g(x) \leq 0\}$

$\Downarrow$

You may need more tools to analyse conditions on boundary

For example, consider the problem of finding the cone with minimum volume that contains a sphere of radius R. (see page 231 of convex opt notes) http://www.cse.iitb.ac.in/~cs717/notes/classNotes/BasicsOfConvexOptimization.pdf

Example problem:

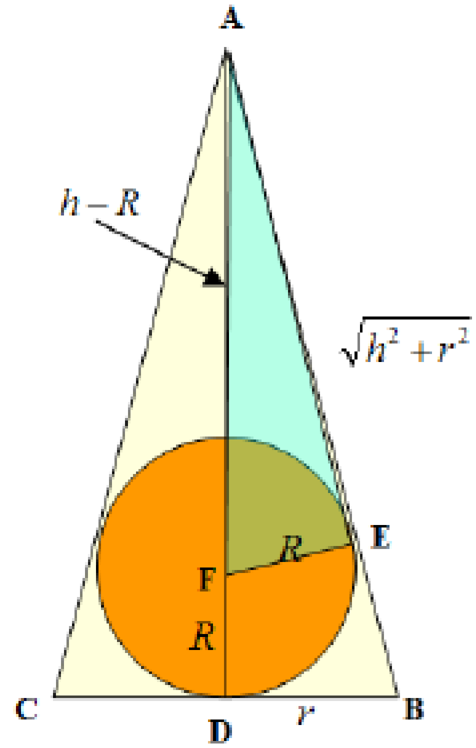


Figure 4.11: Illustrating the constraints for the optimization problem of finding the cone with minimum volume that can contain a sphere of radius R.

Fortunately, this optimization problem (see notes for solution)

1. has only an equality constraint
2. You can substitute for one var in terms of another
   r^2 h
3. Domain is open

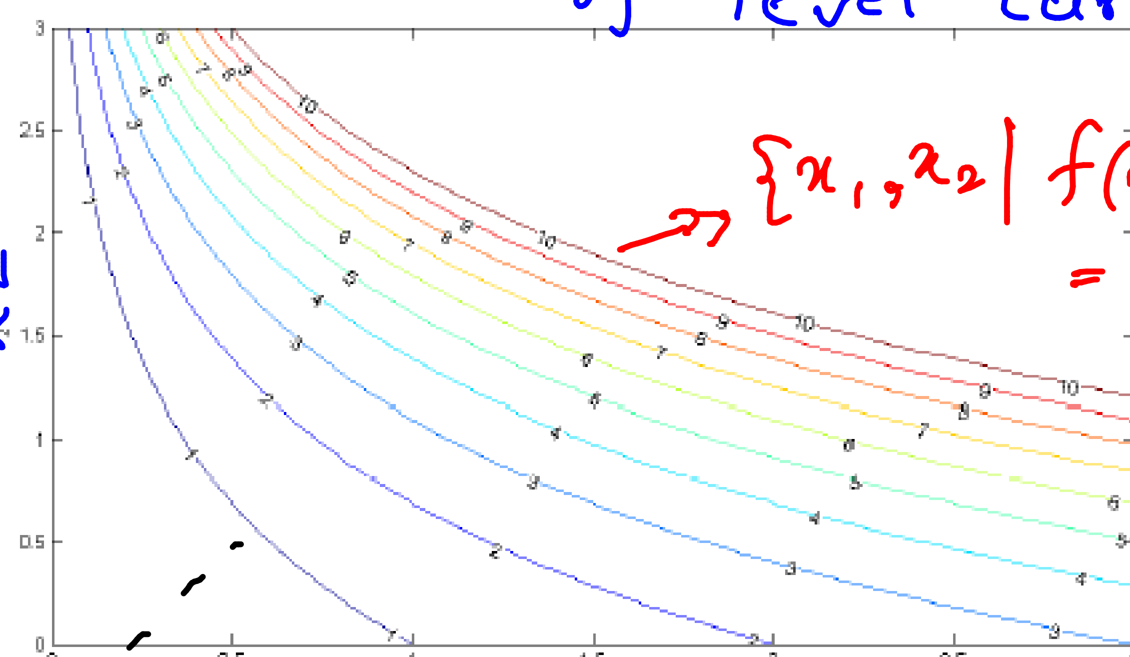
Let us analyse how to solve this problem without making use of these simplifications

Let us go back to

$$\min_{x} \; f(x)$$

$$\text{s.t} \;\; g_1(x) \le 0$$

We will need the concept of level curve



x_2

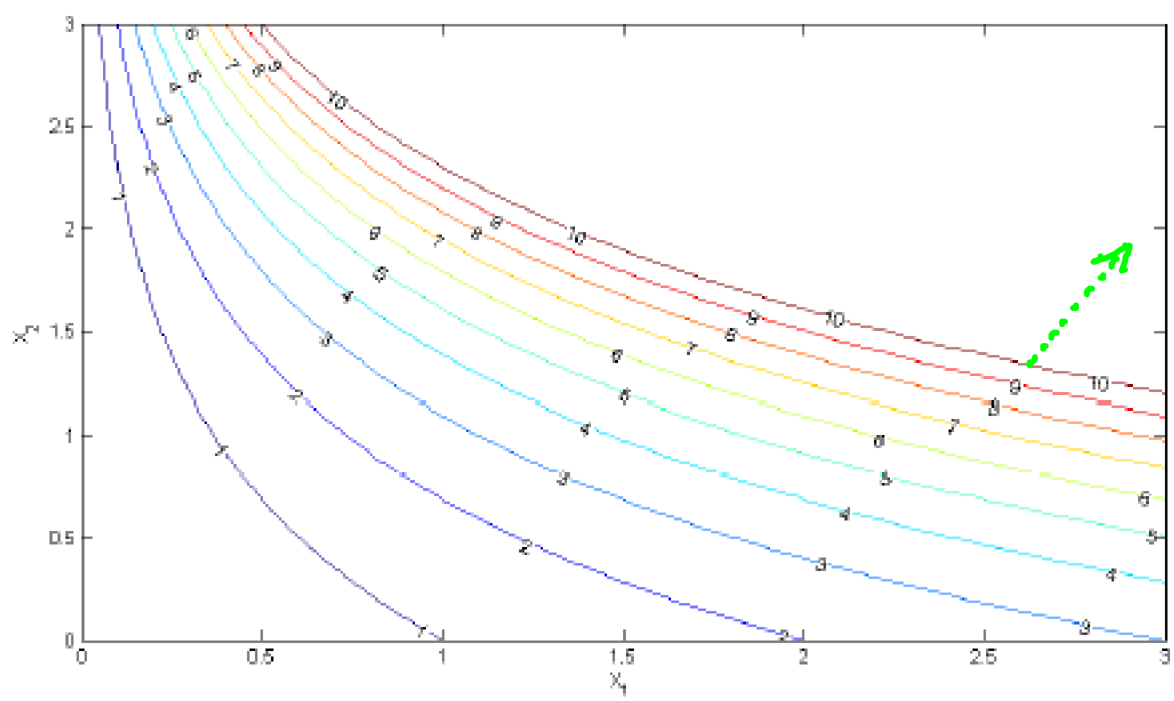
x_1

$\{x_1, x_2 \mid f(x_1, x_2) = 10\}$

$\{x_1, x_2 \mid f(x_1, x_2) = 1\}$

Figure 4.12: 10 level curves for the function $f(x_1, x_2) = x_1 e^{x_2}$.

It turns out that $\nabla f(x_1, x_2)$ is perpendicular to (tangent to) the level curve at (x_1, x_2)

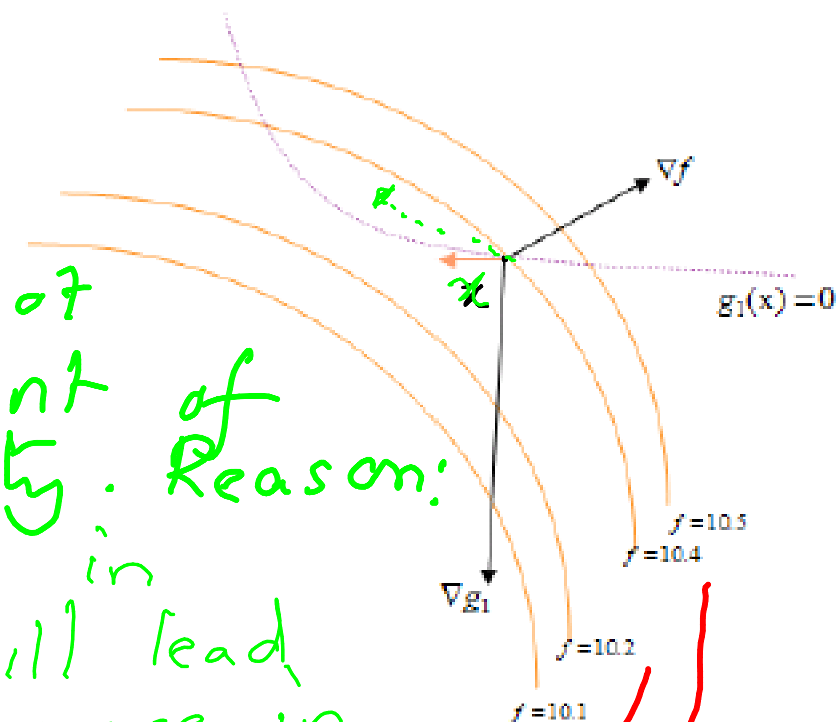


$\nabla f(2.5, 1.25)$

Figure 4.12: 10 level curves for the function $f(x_1, x_2) = x_1 e^{x_2}$.

[For more details & proofs pls look at pg 233 onwards of convex opt notes]

$$\min f(x)$$
$$\text{s.t. } g(x) \le 0$$



level curve for $\{x \mid g(x) = 0\}$

level curves for f

Claim:

x cannot be a point of optimality. Reason: Direction in green will lead to (a) decrease in f since it has a component in dirn of $-\nabla f(x)$ (b) g will also decrease

Let us say x^* is pt of optimality

(1) if $g(x^*) < 0$, $\nabla f(x^*) = 0$

(2) if $g(x^*) = 0$, then we do not want $-\nabla f(x^*)$ to have any component perpendicular to $\nabla g(x^*)$

proportional to

i.e $\nabla f(x^*) \propto \nabla g(x^*)$

Further, they should be in opposite directions

i.e $\nabla f(x^*) = -\lambda \nabla g(x^*)$

Necessary condition

Putting together these 2 conditions: (Necessary conditions)

$$\nabla f + \lambda \nabla g(x^*) = 0 \rightarrow (a)$$
$$\lambda g(x^*) = 0 \rightarrow (b)$$
$$g(x^*) \leq 0 \rightarrow (c)$$

Q: what abt multiple constraints as in SVM

$$\min f(x)$$

$$\text{st} \quad g_1(x) \leq 0$$

$$\vdots$$

$$g_m(x) \leq 0$$

Claim: If x^* (pt of optimality) lies on atleast one of the level curves

$$g_k(x^*) = 0 \quad \text{for } k \in [1..m]$$

then

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) = 0$$

i.e, $\nabla f(x^*)$ should lie in the space spanned by the $\nabla g_i(x^*)$ with special conditions $\lambda_i \geq 0$

So necessary conditions for

$$\min f(x)$$
$$s.t \quad g_1(x) \le 0$$
$$\vdots$$
$$g_m(x) \le 0$$

ARE:

gradient of Lagrange function $= 0$

1. $\nabla f(x^*) + \sum_i \lambda_i \nabla g(x^*) = 0$

$\forall i \in [1, m]$

2. $\lambda_i g_i(x) = 0$ (complementary slackness)

3. $\lambda_i \geq 0$ (λ_i's are lagrange multipliers)

4. & $g_i(x) \leq 0$

H/W: Apply these

Conditions (necessary) to SVM (soft SVM) objective.