Math, Problem Set #5, Convex Analysis

Elysa Strunin

July 21, 2017

7.1

Prove: If S is a nonempty subset of V, then conv(S) is convex.
S is a nonempty subset of V
conv(S) is the set of all finite sums of the form
$\lambda_1\mathbf{x_1} + ... + \lambda_k\mathbf{x_k}, \mathbf{x_i} \in S, k \in N$
where $\lambda_i \geq 0$ and $\lambda_1 + ... + \lambda_k = 1$
Consider the subset of all sums of 2-element combinations:
i.e., $\lambda_1\mathbf{x_1} + ... + \lambda_k\mathbf{x_k}$ where $\lambda_n = 0$ for all but 2 of the elements, "x_i" and "x_j"
$\Rightarrow$ Given λ_i, then $\lambda_j = 1 - \lambda_i \Rightarrow$ conv(S) is convex

7.2

i

Take hyperplane P in V, where $P = \{\mathbf{x} \in V | \langle \mathbf{a}, \mathbf{x} \rangle = b\}$
$\forall \mathbf{x}, \mathbf{y} \in V$, $\langle \mathbf{a}, \mathbf{x} \rangle = b$ and $\langle \mathbf{a}, \mathbf{y} \rangle = b$
Show: $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in V \; \forall \; 0 \leq \lambda \leq 1$
$\langle \mathbf{a}, \lambda\mathbf{x} + (1-\lambda)\mathbf{y} \rangle$
$= \langle \mathbf{a}, \lambda\mathbf{x} \rangle + \langle \mathbf{a}, \mathbf{y} \rangle + \langle \mathbf{a}, -\lambda\mathbf{y} \rangle$
$= \lambda\langle \mathbf{a}, \mathbf{x} \rangle + b - \lambda\langle \mathbf{a}, \mathbf{x} \rangle$
$= \lambda b + b - \lambda b = b$

ii

The proof of convexity for a half space is analogous; just replace "=" with "$\leq$"

7.4

For nonempty, convex, closed C:
$\mathbf{p} \in C$ is $proj_C\mathbf{x}$, if $||\mathbf{x} - \mathbf{p}|| \leq ||\mathbf{x} - \mathbf{y}|| \; \forall \mathbf{y} \in C$

i

Rewritten as all inner products:
Show: $\langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle = \langle \mathbf{x} - \mathbf{p}, \mathbf{x} - \mathbf{p} \rangle + \langle \mathbf{p} - \mathbf{y}, \mathbf{p} - \mathbf{y} \rangle + 2\langle \mathbf{x} - \mathbf{p}, \mathbf{p} - \mathbf{y} \rangle$
RHS $= \langle \mathbf{x}, \mathbf{x} \rangle - 2\langle \mathbf{x}, \mathbf{p} \rangle + 2\langle \mathbf{p}, \mathbf{p} \rangle - 2\langle \mathbf{p}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle + 2\langle \mathbf{x}, \mathbf{p} \rangle - 2\langle \mathbf{x}, \mathbf{y} \rangle - 2\langle \mathbf{p}, \mathbf{p} \rangle + 2\langle \mathbf{p}, \mathbf{y} \rangle$
$= \langle \mathbf{x}, \mathbf{x} \rangle - 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} = \langle \mathbf{x} - \mathbf{y}, \mathbf{x} - \mathbf{y} \rangle$

ii

If (7.14) holds, then all RHS terms of the below equation are ≥ 0.
(All squared norms are ≥ 0).
$||\mathbf{x}-\mathbf{y}||^2 = ||\mathbf{x}-\mathbf{p}||^2 + ||\mathbf{p}-\mathbf{y}||^2 + 2\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y}\rangle$
$\Rightarrow ||\mathbf{x}-\mathbf{y}||^2 > ||\mathbf{x}-\mathbf{p}||^2 \ \forall \mathbf{y} \neq \mathbf{p}$
Take square root: $||\mathbf{x}-\mathbf{y}|| > ||\mathbf{x}-\mathbf{p}||$

iii

Use (i) to write $||\mathbf{x}-\mathbf{z}||^2 = ||\mathbf{x}-\mathbf{p}||^2 + ||\mathbf{p}-\lambda\mathbf{y}-\mathbf{p}+\lambda\mathbf{p}||^2 + 2\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\lambda\mathbf{y}-\mathbf{p}+\lambda\mathbf{p}\rangle$
$= ||\mathbf{x}-\mathbf{p}||^2 + ||\lambda(\mathbf{p}-\mathbf{y})||^2 + 2\langle \mathbf{x}-\mathbf{p}, \lambda(\mathbf{p}-\mathbf{y})\rangle$
$= ||\mathbf{x}-\mathbf{p}||^2 + \lambda^2||\mathbf{p}-\mathbf{y}||^2 + 2\lambda\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y}\rangle$

iv

Have $||\mathbf{x}-\mathbf{z}|| \geq ||\mathbf{x}-\mathbf{p}||$ from definition of projection
Rewrite (7.15): $||\mathbf{x}-\mathbf{z}||^2 - ||\mathbf{x}-\mathbf{p}||^2 = 2\lambda\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y}\rangle + \lambda^2||\mathbf{y}-\mathbf{p}||^2$
$||\mathbf{x}-\mathbf{z}||^2 - ||\mathbf{x}-\mathbf{p}||^2 \geq 0$ because definition of projection
$\Rightarrow 2\lambda\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y}\rangle + \lambda^2||\mathbf{y}-\mathbf{p}||^2 \geq 0$
Divide by λ: $2\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y}\rangle + \lambda||\mathbf{y}-\mathbf{p}||^2 \geq 0$

Overall proof

If $\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y}\rangle \geq 0$, then $\mathbf{p} \in C$ is $proj_C\mathbf{x}$:
(7.14) and (i) $\Rightarrow$ (ii) (definition of projection)
If $\mathbf{p} \in C$ is $proj_C\mathbf{x}$, then $\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y}\rangle \geq 0$:
(7.15) $\Rightarrow$ (iv)
Consider $\lambda = 0$; Clear that $\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y}\rangle \geq 0$ to maintain the inequality in (iv)

7.6

$f : R^n \longrightarrow R$ is a convex function $\Rightarrow \forall \mathbf{x_1}, \mathbf{x_2} \in R^n$:
$f(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \leq \lambda f(\mathbf{x_1}) + (1-\lambda)f(\mathbf{x_2})$
Consider set $J = \{\mathbf{x} \in R^n | f(\mathbf{x}) \leq c\} \subset R^n$
Have $f(\mathbf{x_1}) \leq c, f(\mathbf{x_2}) \leq c \ \forall \mathbf{x_1}, \mathbf{x_2} \in J$
Substitute these bounds into the convex function inequality:
$f(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \leq \lambda c + (1-\lambda)c = c$
$\Rightarrow \forall \mathbf{x_1}, \mathbf{x_2} \in J,\ \lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2} \in J$
$\Rightarrow J$ is a convex set

7.7

$f(\mathbf{x}) = \sum_{i=1}^k \lambda_i f_i(\mathbf{x})$ is convex under the stated conditions:
Because each f_i is convex, $\forall \mathbf{x_1}, \mathbf{x_2} \in C$ and for $0 \leq \lambda \leq 1$:
$f(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) = \sum_{i=1}^k \lambda_i f_i(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2})$
$\leq \sum_{i=1}^k \lambda_i[\lambda f_i(\mathbf{x_1}) + (1-\lambda)f_i(\mathbf{x_2})] = \lambda[\sum_{i=1}^k \lambda_i f_i(\mathbf{x_1})] + (1-\lambda)[\sum_{i=1}^k \lambda_i f_i(\mathbf{x_2})]$
$\Rightarrow f$ is convex

7.13

If $f : R^n \longrightarrow R$ is convex and bounded above, f is constant:
Suppose that f is not constant
i.e., $\exists \mathbf{x}, \mathbf{y}$ s.t. $f(\mathbf{x}) < f(\mathbf{y})$
Function $g(t) = f(\mathbf{x} + t(\mathbf{y} - \mathbf{x}))$ is convex, with $f(\mathbf{x}) = g(0) < f(\mathbf{y}) = g(1)$
Jensen's Inequality: $g(1) \leq \frac{t-1}{t} g(0) + \frac{1}{t} g(t)\ \forall t > 1$
$\Rightarrow g(t) \geq tg(1) - (t-1)g(0) = g(0) + t(g(1) - g(0))$
$\Rightarrow g$ increases unbounded as $t \longrightarrow \infty$,
which contradicts boundedness of f
$\Rightarrow f$ is constant

7.20

Convex $\Rightarrow \forall \mathbf{x_1}, \mathbf{x_2} \in R^n$ and $0 \leq \lambda \leq 1$,
Have $f(\lambda \mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \leq \lambda f(\mathbf{x_1}) + (1-\lambda) f(\mathbf{x_2})$
and $-f(\lambda \mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \leq -\lambda f(\mathbf{x_1}) + (\lambda - 1) f(\mathbf{x_2})$
$\Rightarrow f(\lambda \mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \geq \lambda f(\mathbf{x_1}) + (1-\lambda) f(\mathbf{x_2})$
$\Rightarrow f(\lambda \mathbf{x_1} + (1-\lambda)\mathbf{x_2}) = \lambda f(\mathbf{x_1}) + (1-\lambda) f(\mathbf{x_2})$
$\Rightarrow f$ is affine

7.21

If $\mathbf{x}*$ is a minimizer of $f \Rightarrow \mathbf{x}*$ is a minimizer of $\phi \circ f(\mathbf{x}*)$:
$\mathbf{x}*$ is a minimizer of $f \Rightarrow Df(\mathbf{x}*)(\mathbf{y} - \mathbf{x}*) \geq 0, \forall \mathbf{y} \in \Omega$
ϕ is strictly increasing $\Rightarrow D\phi(f(\mathbf{x}*)) > 0$
So $D\phi(f(\mathbf{x}*))Df(\mathbf{x}*)(\mathbf{y} - \mathbf{x}*) \geq 0, \forall \mathbf{y} \in \Omega$
$\Rightarrow \mathbf{x}*$ is a minimizer of $\phi \circ f(\mathbf{x}*)$

If $\mathbf{x}*$ is a minimizer of $\phi \circ f(\mathbf{x}*) \Rightarrow \mathbf{x}*$ is a minimizer of f:
$\mathbf{x}*$ is a minimizer of $\phi \circ f(\mathbf{x}*) \Rightarrow D\phi(f(\mathbf{x}*))Df(\mathbf{x}*)(\mathbf{y} - \mathbf{x}*) \geq 0, \forall \mathbf{y} \in \Omega$
ϕ is strictly increasing $\Rightarrow D\phi(f(\mathbf{x}*)) > 0$
So $Df(\mathbf{x}*)(\mathbf{y} - \mathbf{x}*) \geq 0, \forall \mathbf{y} \in \Omega$
$\Rightarrow \mathbf{x}*$ is a minimizer of f