For $y \in \mathbb{R}^n$ with $x+y \in S$ there exists $x' \in [x, x+y]$ so that

$$f(x+y) = f(x) + \nabla f(x)^T y + \frac{1}{2} y^T Hf(x') y$$

By assumption c) and by Proposition 3 we have

$$y^T Hf(x') y \geq m\, y^T y.$$

Hence,

$$f(x+y) \geq f(x) + \nabla f(x)^T y + \frac{1}{2} m \|y\|^2.$$

The right hand side is a convex quadratic function (in y) with gradient $\nabla f(x) + my$. Its minimum is attained at $y^* = -\frac{1}{m} \nabla f(x)$. Choosing y so that $f(x+y) = p^*$ yields

$$p^* = f(x+y) \geq f(x) + \nabla f(x)^T\left(-\frac{1}{m} \nabla f(x)\right) + \frac{1}{2} m \left(\frac{1}{m} \nabla f(x)\right)^2$$

$$= f(x) - \frac{1}{2m} \|\nabla f(x)\|^2. \qquad \boxtimes$$

<u>Lemma 5</u> Under the assumptions above we have for all $i \geq 1$

$$\frac{L}{2m^2} \|\nabla f(x^{(i+1)})\| \leq \left(\frac{L}{2m^2} \|\nabla f(x^{(i)})\|\right)^2.$$

<u>Proof</u> Define $x^+ = x^{(i+1)}$,

$$x = x^{(i)}, \quad x_{nt} = -Hf(x)^{-1} \nabla f(x).$$

We have

$$\begin{aligned}
\| \nabla f(x') \| &= \| \nabla f(x + x_{nt}) \overbrace{- \nabla f(x) - Hf(x)\, x_{nt}}^{=0} \| \\
&\overset{\text{chain rule}}{=} \| \int_0^1 (Hf(x + t\, x_{nt}) - Hf(x))\, x_{nt}\, dt \| \\
&\overset{\text{triangle ineq.}}{\leq} \int_0^1 \| (Hf(x + t\, x_{nt}) - Hf(x)) x_{nt} \| \, dt \\
&\overset{\text{matrix vs. vector norm}}{\leq} \int_0^1 \| Hf(x + t\, x_{nt}) - Hf(x) \| \; \| x_{nt} \| \, dt \\
&\overset{\text{assumption d)}}{\leq} \int_0^1 L \, \| t\, x_{nt} \| \; \| x_{nt} \| \, dt \\
&= \frac{L}{2} \| x_{nt} \|^2 \\
&= \frac{L}{2} \| Hf(x)^{-1} \nabla f(x) \|^2 \\
&\overset{\frac{1}{m} \geq \text{ largest eigenvalue of } Hf(x)^{-1}}{\leq} \frac{L}{2m^2} \| \nabla f(x) \|^2 .
\end{aligned}$$

$\boxtimes$

Theorem 6 Algorithm 1 is correct.

Proof Apply Lemma 5 inductively and use assumption e).

For $j \leq i$ we get

$$\frac{L}{2m^2} \| \nabla f(x^{(i)}) \| \leq \Big(\frac{L}{2m^2} \underbrace{\| \nabla f(x^{(j)}) \|}_{\leq \frac{m^2}{L}} \Big)^{2^{j-i}}$$

So

$$\|\nabla f(x^{(N)})\| \leq \frac{2m^2}{L} \left(\frac{1}{2}\right)^{2^N}.$$

Applying Lemma 4 yields

$$f(x^{(N)}) - p^* \leq \frac{1}{2m}\left(\frac{2m^2}{L}\left(\frac{1}{2}\right)^{2^N}\right)^2$$

$$= \frac{2m^3}{L^2}\left(\frac{1}{2}\right)^{2^{N+1}}. \qquad \boxtimes$$

Chapter III Conic optimisation

§ 1 Convex cones

Def. 1 a) A nonempty set $K \subseteq \mathbb{R}^n$ is called <u>(convex) cone</u> if $\forall \alpha, \beta \in \mathbb{R}_{\geq 0}\ \forall x, y \in K$:

$$\alpha x + \beta y \in K.$$

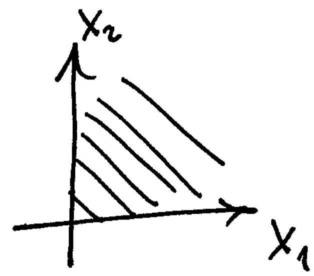
b) A convex cone K is called <u>pointed</u> if $K \cap (-K) = \{0\}$.

c) A convex cone K is called <u>proper</u> if K is pointed, closed and full dimensional (i.e. $\operatorname{int} K \neq \emptyset$).

Examples 2

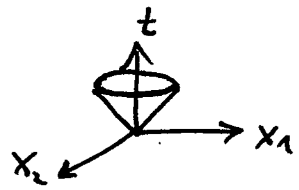
a) <u>nonnegative orthant</u>

$$\mathbb{R}^n_{\geq 0} = \{x \in \mathbb{R}^n : x_1 \geq 0, \ldots, x_n \geq 0\}$$
$$= \mathbb{R}_{\geq 0} \times \cdots \times \mathbb{R}_{\geq 0}$$
$$= \operatorname{cone}\{e_1, \ldots, e_n\}$$



b) <u>Lorentz cone</u> $\hat{=}$ ice cream cone

$$\mathcal{L}^{n+1} = \{(x, t) \in \mathbb{R}^{n+1} : \|x\| = \sqrt{\textstyle\sum_i x_i^2} \leq t\}$$



c) Cone of positive semidefinite matrices

$\hat{=}$ semidefinite cone $\hat{=}$ PSD cone

$S^n = \{ X \in \mathbb{R}^{n\times n} : X \text{ symmetric}, X_{ij} = X_{ji} \}$

vector space of symmetric matrices, $\dim S^n = \frac{n(n+1)}{2}$.

$S^n_{\succeq 0} = \{ X \in S^n : X \text{ is positive semidefinite} \}$.

$x^T X x \geq 0$ for all $x \in \mathbb{R}^n$.

Proper cones are positivity domains. They define a partial order on $\mathbb{R}^n$:

$x \geq_K y \iff x - y \in K$.

Familiar example: $K = [0, \infty) \subseteq \mathbb{R}$. Then $x \geq_K y$ iff $x \geq y$.

Def. 3 Let $K \subseteq \mathbb{R}^n$ be a convex cone. Its dual cone is defined as $K^* = \{ y \in \mathbb{R}^n : x^T y \geq 0 \text{ for all } x \in K \}$.

Proposition 4 Let $K \subseteq \mathbb{R}^n$ be a convex cone. Then,

a) $K \subseteq (K^*)^*$

b) If K is closed, then $K = (K^*)^*$.