Overview

- We can find Δx as the change in x along some steepest descent direction of f without constraints
- Thus, let $x_u^{k+1} = x^k + \Delta x$ be the working set that reduces $f(x)$ without constraints (unbounded)
- To find the constrained working set, we project x_u^{k+1} onto Ω to get x^{k+1}

$$\Omega = \bigcap_{i=1}^{m} \{x \mid g_i(x) \leq 0\}$$

Algo: Initialise: $x_u^{(0)}$, $x_p^{(0)} = P_\Omega(x_u^{(0)})$

until convergence, $x_u^{(k+1)} = x_u^{(k)} - t^{(k)} \nabla f(x^{(k)})$

$x_p^{(k+1)} = P_\Omega(x_u^{(k+1)})$

- To project x_u onto the non-empty closed convex set Ω to get the projected point x_p, we solve:

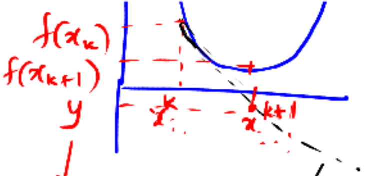
$$x_p = P_\Omega(x_u) = \operatorname*{argmin}_{z \in \Omega} \|x_u - z\|_2^2$$

- That is, the projected point x_p is the point in Ω that is the closest to the unbounded optimal point x_u if Ω is a non-empty closed convex set

Recall: If g_i's are lower semi-cts then Ω is closed convex & a unique x_p is guaranteed to exist

Descent direction for a convex function

- For a descent in a convex function f, we must have $f(x^{k+1}) \geq$ Value at x^{k+1} obtained by linear interpolation from x^k
- *ie.* $f(x^{k+1}) \geq f(x^k) + \nabla^\top f(x^k)(x^{k+1} - x^k)$



$$y = f(x^k) + \nabla^T f(x^k)(x^{k+1} - x^k)$$

- Thus, for Δx^k to be a descent direction, it is necessary that $\nabla^\top f(x^k)\Delta x^k \leq 0$ (where $\Delta x^k = x^{k+1} - x^k$)

Q: is $\nabla^T f(x^k)\left(P_\Omega\left(x^k - t^k \nabla f(x^k)\right) - x^k\right) \leq 0$?

We want that the point obtained after the projection of x_u^{k+1} to be a descent from x^k for the function f

$$\nabla f(x^k) \cdot \Delta x_p \leq 0$$

This is only necessary ... for complete convergence

(where $\Delta x_p = P_\Omega(x_u^{k+1}) - x^k$)

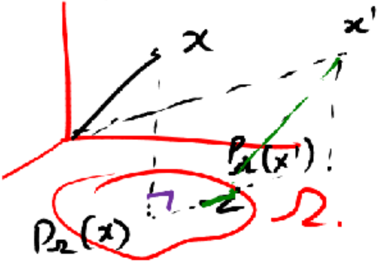
Ref Nemirovski Sec 5.3.1

- **Claim:** If $P_\Omega(x)$ is a projection of x, then

} Actually iff claim

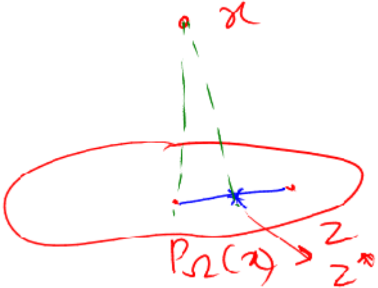
$$\left(z - P_\Omega(x)\right)^\top \left(x - P_\Omega(x)\right) \leq 0, \ \forall z \in \Omega$$

- That is, the angle between $\left(z - P_\Omega(x)\right)$ and $\left(x - P_\Omega(x)\right)$ is obtuse (or right-angled for the projected point), $\forall z \in \Omega$



Proof for $\langle z - P_\Omega(x), x - P_\Omega(x) \rangle \le 0$

- To be more general, let us consider an inner product $\langle a, b \rangle$ instead of $a^\top b$
- Let $z^* = (1-\alpha)P_\Omega(x) + \alpha z$, for some $\alpha \in (0,1)$, and $z \in \Omega$ $\implies z^* = P_\Omega(x) + \alpha(z - P_\Omega(x))$, $z^* \in \Omega$



- Since $P_\Omega(x) = \mathrm{argmin}_{z \in \Omega} \|x - z\|_2^2$, $\left\|x - P_\Omega(x)\right\|^2 \le \|x - z^*\|^2$

$$
\begin{aligned}
&\|x - z^*\|^2 \\
&= \left\| x - \left(P_\Omega(x) + \alpha(z - P_\Omega(x))\right) \right\|^2 \\
&= \|x - P_\Omega(x)\|^2 + \alpha^2 \|z - P_\Omega(x)\|^2 \underbrace{- 2\alpha \langle x - P_\Omega(x), z - P_\Omega(x) \rangle}_{\geq 0} \\
&\geq \|x - P_\Omega(x)\|^2
\end{aligned}
$$

$$
\implies \underbrace{\langle x - P_\Omega(x), z - P_\Omega(x) \rangle}_{\text{LHS}} \leq \frac{\alpha}{2} \|z - P_\Omega(x)\|^2, \; \forall \alpha \in (0, 1)
$$

- Thus, the LHS can either be 0 or a negative value. Any positive value of the LHS will lead to a contradiction for some small $\alpha \to 0$
- Hence, we proved that $\langle z - P_\Omega(x), x - P_\Omega(x) \rangle \leq 0$

- We can also prove that if $\langle x - x^*, z - x^* \rangle \leq 0$, $\forall z \in \Omega$ s.t. $z \neq x^*$, and $x^* \in \Omega$, then

$$x^* = P_\Omega(x) = \underset{\bar{z} \in \Omega}{\operatorname{argmin}} \|x - \bar{z}\|_2^2$$

- Consider $\|x - z\|^2 - \|x - x^*\|^2$
$= \left\|x - x^* + (x^* - z)\right\|^2 - \|x - x^*\|^2$
$= \|x - x^*\|^2 + \|z - x^*\|^2 - 2\langle x - x^*, z - x^* \rangle - \|x - x^*\|^2$
$= \|z - x^*\|^2 - 2\langle x - x^*, z - x^* \rangle$
> 0
- $\implies \|x - z\|^2 > \|x - x^*\|^2$, $\forall z \in \Omega$ s.t. $z \neq x^*$
- This proves that $x^* = P_\Omega(x)$

References

- Yu-Hong Dai, Roger Fletcher. New algorithms for singly linearly constrained quadratic programs subject to lower and upper bounds. `http://link.springer.com/content/pdf/10.1007%2Fs10107-005-0595-2.pdf`