$\|\cdot\|$ be norm on $\mathbb{R}^n$

Dual norm $\|\cdot\|_*$ is defined
(conjugate)

$$\|z\|_* = \sup\{ |z^T x| \mid \|x\| \le 1 \}$$

(i) $\|\alpha z\|_* = |\alpha| \|z\|_*$

(ii) $\|z\|_* = 0 \iff z = 0$.

(iii) $\|z_1 + z_2\|_* = \sup\{ |(z_1 + z_2)^T x| \mid \|x\| \le 1 \}$

$$\le \sup\{ |z_1^T x| + |z_2^T x| \mid \|x\| \le 1 \} \le \sup\{ |z_1^T x| \mid \|x\| = 1 \} + \sup\{ |z_2^T x| \mid \|x\| = 1 \}$$

$$= \|z_1\|_* + \|z_2\|_*$$

$$\|z\|_* = \sup \{ |z^T x| \mid \|x\|_p \leq 1 \}.$$

Conjugate norm of p norm.

$$|z^T x| \leq \|z\|_q \|x\|_p \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1$$

$$\|z\|_* = \|z\|_q.$$

Excercise:

[make these arguments more precise]

## Steepest Descent

Initial guess, $x \in \mathrm{dom} f$ , $\epsilon$

do until stopping criterion

(i) Choose $\Delta x_{sd}$

(ii) Line search $t$ (Backtracking line search / Exact line search)

(iii) Update $x \leftarrow x + t \Delta x_{sd}$.

$$\Delta x_{nsd} = \operatorname{Arg\,min} \left\{ \nabla f(x)^T v \;\middle|\; \|v\| \le 1 \right\}$$

$$\Delta x_{sd} = \boxed{\|\nabla f(x)\|_* \, \Delta x_{nsd}}$$

$$\|\cdot\|_* = \|\cdot\|_q \quad \left(\frac{1}{p} + \frac{1}{q} = 1\right)$$

Note

(i) $\boxed{\|\Delta x_{sd}\| = \|\nabla f(x)\|_*}$

(ii) $-\|\nabla f(x)\|_* \|v\| \le \nabla f(x)^T v \le \|\nabla f(x)\|_* \|v\| \qquad \sqrt{x^T P x} \qquad P > 0$

$$-\|\nabla f(x)\|_* \le \nabla f(x)^T \Delta x_{nsd} \le \|\nabla f(x)\|_*$$

(iii) $\nabla f(x)^T \Delta x_{nsd} = -\|\nabla f(x)\|_*$

$$\boxed{\nabla f(x)^T \Delta x_{sd} = -\|\nabla f(x)\|_*^2}$$

# Convergence Analysis of Steepest Descent

$f$ strongly convex on $S$

$\nabla^2 f(x) \preceq MI$

$$f(x + t\Delta x_{sd}) \le f(x) + t\nabla f(x)^T \Delta x_{sd} + \frac{Mt^2}{2}\|\Delta x_{sd}\|_2^2$$

$\|x\| \geq \gamma \|x\|_2 \quad \gamma \in (0,1]$

$\|x\|_* \geq \tilde{\gamma} \|x\|_2 \quad \tilde{\gamma} \in (0,1]$

$\|x\|_q \le \|x\|_p \le n^{\frac{1}{p}-\frac{1}{q}} \|x\|_q$

$p < q$

$$f(x + t\Delta x_{sd}) \leq f(x) + t\nabla f(x)^T \Delta x_{sd} + \frac{M}{2\gamma^2} \|\Delta x_{sd}\|^2 t^2$$

$$\nabla f(x)^T \Delta x_{sd} = - \|\nabla f(x)\|_*^2$$

$$f(x + t\Delta x_{sd}) \leq f(x) - t\|\nabla f(x)\|_*^2 + \frac{Mt^2}{2\gamma^2} \|\Delta x_{sd}\|^2$$

$$f(x + t\Delta x_{sd}) \leq f(x) + \|\nabla f(x)\|_*^2 \left( -t + \frac{Mt^2}{2\gamma^2} \right)$$

Choose $\hat{t} = \underline{\gamma^2/M}$

$$f(x + \hat{t}\,\Delta x_{sd}) \leq f(x) - \frac{\gamma^2}{2M} \|\nabla f(x)\|_*^2$$

Backtracking line search.

$$\alpha < \frac{1}{2}, \quad \boxed{\nabla f(x)^T \Delta x_{sd} = - \|\nabla f(x)\|_*^2}$$

$$\boxed{f(x + \hat{t}\,\Delta x_{sd}) \leq f(x) + \alpha \hat{t}\, \boxed{\nabla f(x)^T \Delta x_{sd}}}$$

$$\boxed{t \geq \min \left\{ 1, \beta \frac{\gamma^2}{M} \right\}}$$

stepsize given by Backtracking line search. follows this condition.

$$f(x + t\,\Delta x_{sd}) = f(x^{(k+1)})$$

$$\leq f(x^{(k)}) - \underbrace{\alpha \min\left\{1, \frac{\beta\gamma^2}{M}\right\}} \|\nabla f(x^{(k)})\|_*^2$$

$$t \geq \min\left\{1, \frac{\beta\gamma^2}{M}\right\}$$

$$f(x^{(k+1)}) \leq f(x^{(k)}) - \min\left\{\alpha, \alpha\beta\frac{\gamma^2}{M}\right\} \|\nabla f(x^{(k)})\|_*^2$$

$$f(x^{(k+1)}) - p^* \leq \underline{f(x^{(k)}) - p^*} - \min\left\{\alpha, \alpha\beta\frac{\gamma^2}{M}\right\} \underline{\|\nabla f(x^{(k)})\|_*^2}$$

$$\|\nabla f(x^{(k)})\|_*^2 \geq \tilde{\gamma}^2 \|\nabla f(x^{(k)})\|_2^2$$

$$f(x^{(k+1)}) - p^* \leq f(x^{(k)}) - p^* - \min\left\{ \alpha \tilde{\gamma}^2, \frac{\alpha\beta \gamma^2 \tilde{\gamma}^2}{M} \right\} \underline{\|\nabla f(x^{(k)})\|_2^2}$$

$$f(x^{(k)}) - p^* \leq \frac{1}{2m} \|\nabla f(x^{(k)})\|_2^2$$

$$f(x^{(k+1)}) - p^* \leq \underbrace{\left(1 - \min\left\{ \alpha \tilde{\gamma}^2 m, \alpha\beta (\gamma\tilde{\gamma})^2 \frac{m}{M} \right\}\right)}_{c} \left(f(x^{(k)}) - p^*\right)$$

$$\left( f(x^{(k)}) - p^* \right) \leq c^k \left( f(x^{(0)}) - p^* \right)$$

$$c = 1 - \min \left\{ \alpha \tilde{\gamma}^2 m \, , \; \alpha \beta (\gamma \tilde{\gamma})^2 \frac{m}{M} \right\} < 1$$

$c^k \to 0$ as $k \to \infty$.

$f(x^{(k)}) - p^* \longrightarrow 0$ as $k \to \infty$.