3]

$$\min_{\|d\|_P=1} d^T \nabla f(x)$$

$$h(d,\lambda) = d^T \nabla f(x) + \lambda\left(\|d\|_P^2 - 1\right)$$

$$\nabla_1 h(d,\lambda) = \nabla f(x) + 2\lambda P d \stackrel{!}{=} 0$$

$$\Rightarrow \quad d = -\frac{1}{2\lambda} P^{-1} \nabla f(x)$$

$$\|d\|_P \stackrel{!}{=} 1 \quad \Rightarrow \quad d = -\frac{P^{-1}\nabla f(x)}{\|P^{-1}\nabla f(x)\|_P}$$

$$f(x) = \frac{1}{2} x^T A x + b^T x$$

$$x^{k+1} = x^k + \alpha_k d^k \quad , \quad d^k = -P^{-1} \nabla f(x^k)$$

$$\alpha_k = \operatorname*{argmin}_{\sigma \geq 0} f(x^k + \sigma d^k)$$

$$= - \frac{(d^k)^T \nabla f(x^k)}{(d^k)^T A d^k} > 0$$

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

$$f(x^{k+1}) - f(\bar{x}) = \underbrace{(\text{----})}_{\leq \cdots \text{ with Kantorovich}} (f(x^k) - f(\bar{x}))$$

$$f(x^{k+1}) = f(x^k + \alpha_k d^k)$$

$$\overset{\text{Taylor}}{=} f(x^k) + \alpha_k \nabla f(x^k)^T d^k + \frac{\alpha_k^2}{2} (d^k)^T A d^k$$

$$= f(x^k) - \frac{1}{2} \frac{(\nabla f(x^k)^T P^{-1} \nabla f(x^k))^2}{\nabla f(x^k)^T P^{-1} A P^{-1} \nabla f(x^k)}$$

$$= f(x^k) - \frac{1}{2} \| P^{-1/2} \nabla f(x^k) \|_2^4 / \| P^{-1/2} \nabla f(x^k) \|^2_{P^{-1/2} A P^{-1/2}}$$

$$P = U \underbrace{\Lambda}_{= \mathrm{diag}(\lambda_1, \dots, \lambda_n)} U^T$$

$$P^\alpha = U \underbrace{\Lambda^\alpha}_{= \mathrm{diag}(\lambda_1^\alpha, \dots, \lambda_n^\alpha)} U^T$$

$$f(x^k) - f(\bar{x}) = \cancel{f(\bar{x})} + \overbrace{\nabla f(\bar{x})}^{= 0 \; = A\bar{x} + b}{}^T (x^k - \bar{x})$$

$$+ \frac{1}{2} (x^k - \bar{x})^T A (x^k - \bar{x})$$

$$- \cancel{f(\bar{x})}$$

$$= \frac{1}{2} (x^k - \bar{x})^T \underbrace{A (x^k - \bar{x})}_{= Ax^k + b = \nabla f(x^k)}$$

$$\Rightarrow f(x^k) - f(\bar{x}) = \frac{1}{2} \nabla f(x^k)^T A^{-1} \nabla f(x^k)$$

$$= \frac{1}{2} \| P^{-1/2} \nabla f(x^k) \|^2_{P^{1/2} A^{-1} P^{1/2}}$$

$$\Rightarrow f(x^{k+1}) - f(\bar{x}) = \left(1 - \frac{\| P^{-1/2} \nabla f(x^k) \|_2^4}{\left(\| P^{-1/2} \nabla f(x^k) \|^2_{P^{-1/2} A P^{-1/2}} \cdot \| P^{-1/2} \nabla f(x^k) \|^2_{P^{1/2} A^{-1} P^{1/2}} \right)} \right) \cdot \left(f(x^k) - f(\bar{x}) \right)$$

Use Kantorovich, $M = P^{-1/2} A P^{-1/2}$

$$\hookrightarrow f(x^{k+1}) - f(\bar{x}) \leq \left(\frac{\Lambda - \lambda}{\Lambda + \lambda}\right)^2 \left(f(x^k) - f(\bar{x})\right)$$

with $\Lambda = \lambda_{max}\left(P^{-1/2} A P^{-1/2}\right)$

$\lambda = \lambda_{min}\left(P^{-1/2} A P^{-1/2}\right)$

$$\Rightarrow \underbrace{f(x^{k+1}) - f(\bar{x})}_{= \frac{1}{2}(x^{k+1} - \bar{x})^T A (x^{k+1} - \bar{x})} \leq \left(\frac{\Lambda - \lambda}{\Lambda + \lambda}\right)^{2(k+1)} \left(f(x^0) - f(\bar{x})\right)$$

$$\|x^k - \bar{x}\|_2^2 \leq \frac{2}{\lambda_{min}(A)}\left(f(x^k) - f(\bar{x})\right) \leq \frac{2}{\lambda_{min}(A)} \left(\frac{\Lambda - \lambda}{\Lambda + \lambda}\right)^{2k} \left(f(x^0) - f(\bar{x})\right)$$

$$\Rightarrow \quad \|x^k - \bar{y}\|_2^2 \leq \frac{\lambda_{max}(A)}{\lambda_{min}(A)} \left(\frac{L-\lambda}{L+\lambda} \right)^{2k} \|x^0 - \bar{x}\|_2^2$$

2] "<="

$$L(\bar{x}, \bar{\lambda}, \bar{\mu}) = f(\bar{x}) + \underbrace{\bar{\lambda}}_{\geq 0}{}^T \underbrace{g(\bar{x})}_{\leq 0} + \bar{\mu}^T \underbrace{h(\bar{x})}_{=0}$$

$$\leq f(\bar{x}) = d(\bar{\lambda}, \bar{\mu}) = \inf_x L(x, \bar{\lambda}, \bar{\mu})$$

$$\leq L(x, \bar{\lambda}, \bar{\mu}) \quad \forall x$$

$$L(\bar{x},\bar{\lambda},\bar{\mu}) = \inf_{x} L(x,\bar{\lambda},\bar{\mu})$$

$$\Rightarrow \quad \bar{\lambda}^T g(\bar{x}) = 0$$

$$L(\bar{x},\bar{\lambda},\bar{\mu}) \geq f(\bar{x}) \underset{\uparrow}{\geq} L(\bar{x},\lambda,\mu) \ \forall (\lambda,\mu)$$

$$g(\bar{x}) \leq 0, \ h(\bar{x}) = 0$$

" $\Rightarrow$ "

$$L(\bar{x},\bar{\lambda},\bar{\mu}) \leq L(x,\bar{\lambda},\bar{\mu}) \quad \forall x$$

$$\Rightarrow \quad d(\bar{\lambda},\bar{\mu}) = L(\bar{x},\bar{\lambda},\bar{\mu}) \geq L(\bar{x},\lambda,\mu) \quad \forall (\lambda,\mu)$$

$$\geq d(\lambda,\mu)$$

$(\bar{\lambda}, \bar{\mu})$ is max. of (D)!

$$
\begin{aligned}
h(\bar{x}, \bar{\lambda}, \bar{\mu}) &= \inf_x h(x, \bar{\lambda}, \bar{\mu}) \\
&\leq \sup_{\lambda, \mu} \inf_x h(x, \lambda, \mu) \\
&\leq \inf_x \sup_{\lambda, \mu} h(x, \lambda, \mu) \\
&\leq \sup_{\lambda, \mu} h(\bar{x}, \lambda, \mu) \leq h(\bar{x}, \bar{\lambda}, \bar{\mu})
\end{aligned}
$$

$$L(\bar{x}, \bar{\lambda}, \mu) = \inf_{x} L(x, \bar{\lambda}, \bar{\mu}) = d(\bar{\lambda}, \bar{\mu})$$

$$= \sup_{\lambda, \mu} L(\bar{x}, \lambda, \mu) = f(\bar{x})$$

$\Rightarrow$ $\bar{x}$ is feasible

weak duality: $f(\tilde{x}) \geq d(\tilde{\lambda}, \tilde{\mu})$

$\Rightarrow$ $\bar{x}$ global min, $(\bar{\lambda}, \bar{\mu})$ global max.