- grad desc $O(K \log(1/\varepsilon))$
  conj grad $O(\sqrt{K} \log(1/\varepsilon))$
  (agm)

$$L_\rho = f(x) + g(z) + \frac{\rho}{2}\|Ax + Bz - c + w\|_2^2 - \frac{\rho}{2}\|w\|_2^2$$

$$L_\rho = f(x) + g(z) + \frac{\rho}{2}\|x - z + w\|_2^2 - \frac{\rho}{2}\|w\|_2^2$$

x: $\quad x^+ = \operatorname{argmin} f(x) + \frac{\rho}{2}\|z - w - x\|_2^2$

$\quad\quad = \operatorname{prox}_{f,1/\rho}(z - w)$

$$L_\rho = \frac{1}{2}\|y - X\beta\|_2^2 + \lambda\|\alpha\|_1 + \frac{\rho}{2}\|\beta - \alpha + w\|_2^2 - \frac{\rho}{2}\|w\|_2^2$$

β: $\quad \beta^+ = \operatorname{argmin} \frac{1}{2}\|y - X\beta\|_2^2 + \frac{\rho}{2}\|\alpha - w - \beta\|_2^2$

$\quad\quad = (X^T X + \rho I)^{-1}(X^T y + \rho(\alpha - w))$

α: $\quad \alpha^+ = \operatorname{argmin} \frac{1}{2}\|\beta + w - \alpha\|_2^2 + \frac{\lambda}{\rho}\|\alpha\|_1$

$\quad\quad = S_{\lambda/\rho}(\beta^+ + w)$

$w^+ = w + \beta^+ - \alpha^+.$

$$\left(\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6 \ldots\right)$$

$$\min_{P} \quad \|X - XP\|_F^2 \quad \text{s.t.} \quad P \text{ being projection matrix of rank } k$$

$$\Longleftrightarrow \max \quad \langle \underbrace{X^T X}_{S}, P \rangle \quad \text{st} \quad ''$$

↑ expand $\|\cdot\|_F$

$$'' \quad \mathrm{tr}(SP)$$

$$\Longleftrightarrow \max \quad \mathrm{tr}(SP) \quad \text{st} \quad P \in \mathcal{F}_k$$

↑ Fan (1949)

$\mathcal{F}_k$ ↓ convex hull of rank k proj. matrices

try taking $x_1 = x_2$, $x_2 = x_3 \ldots$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_B \end{pmatrix} - \begin{pmatrix} x \\ x \\ \vdots \\ x \end{pmatrix} = 0$$

$$L_\rho = \sum f_i(x_i) + \frac{\rho}{2} \left\| \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_B \end{pmatrix} - \begin{pmatrix} x \\ x \\ \vdots \\ x \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_B \end{pmatrix} \right\|_2^2$$

$$= \sum f_i(x_i) + \frac{\rho}{2} \sum \|x_i - x + w_i\|_2^2$$

$$= \sum \left(f_i(x_i) + \frac{\rho}{2} \|x_i - x + w_i\|_2^2 \right)$$