$$w = u/\rho, \text{ i.e. } \rho w = u$$

$$\left.\begin{array}{l} u^T Ax + \frac{\rho}{2}\|Ax-b\|_2^2 \\ \\ \frac{\rho}{2}\|Ax-b+w\|_2^2 \end{array}\right\}$$

same terms involving $x$

$$= \frac{\rho}{2}\|Ax-b\|_2^2 \checkmark + \frac{\rho}{2}\cancel{\|w\|_2^2}$$

$$+ \underbrace{\rho w^T(Ax}_{= u^T Ax} - \cancel{b})$$

---

$$u^+ = u + \rho(Ax-b)$$

$$w^+ = w + Ax - b.$$

---

$$x_1^+ = \underset{x_1}{\operatorname{argmin}}\ f_1(x_1) + \frac{\rho}{2}\|\overbrace{Ax}^{= A_1x_1 + A_2x_2} - b + w\|_2^2$$

$$x_2^+ = \underset{x_2}{\operatorname{argmin}}\ f_2(x_2) + \frac{\rho}{2}\|\overbrace{Ax}^{= A_1x_1^+ + A_2x_2} - b + w\|_2^2$$

$$w^+ = w + Ax - b.$$

---

$$\min\ \frac{1}{2}\|y - X\beta\|_2^2 + \lambda\|\alpha\|_1 \quad \text{s.t. } \alpha = D\beta$$

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

$$= \underset{\beta}{\operatorname{argmin}}\ \frac{1}{2}\|y - X\beta\|_2^2 + \frac{1}{2}\|(\alpha - w)\sqrt{\rho} - \sqrt{\rho}\cdot D\beta\|_2^2$$

$$= (X^TX + \rho D^TD)^{-1}(X^Ty + \rho D^T(\alpha - w))$$

$$\alpha^+ = \underset{\alpha}{\operatorname{argmin}} \; \lambda \|\alpha\|_1 + \frac{\rho}{2} \| D\beta^+ - \alpha + w \|_2^2$$

$$= \underset{\alpha}{\operatorname{argmin}} \; \frac{1}{2} \| D\beta^+ + w - \alpha \|_2^2 + \frac{\lambda}{\rho} \|\alpha\|_1$$

$$= S_{\frac{\lambda}{\rho}} ( D\beta^+ + w )$$

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

---

Implementation tip: factorize $X^TX + \rho D^TD$ $\underline{\underline{\text{once}}}$ (eg. Cholesky), then use it for all $\beta$-updates.

---

$$\min \; \sum f_i(x_i) \quad \text{s.t.} \quad x_i = x, \; i = 1 \dots B \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_B \end{pmatrix} - \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix} = 0$$

$$L(x_1, \dots x_B, x) = \sum_{i=1}^{B} f_i(x_i) + \underbrace{\frac{\rho}{2} \sum_{i=1}^{B} \| x_i - x + w_i \|_2^2}_{\frac{\rho}{2} \left\| \begin{pmatrix} x_1 \\ \vdots \\ x_B \end{pmatrix} - \begin{pmatrix} x \\ \vdots \\ x \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_B \end{pmatrix} \right\|_2^2}$$

---

$$\min \; \sum_{i=1}^{B} f_i(a_i^T x_i + b_i) + g(x)$$

$$\text{s.t.} \quad x_i = x$$

$$f(y) - f(x) = g(y) - g(x) + \Sigma\left(h_i(y_i) - h_i(x_i)\right)$$

$$\geq \nabla g(x)^T (y-x) + \Sigma\, h_i(y_i) - h_i(x_i)$$

$$= \Sigma \underbrace{\left( \frac{\partial g}{\partial x_i}(x)\,(y_i - x_i) + h_i(y_i) - h_i(x_i) \right)}_{\geq 0.}$$

Property : $g(x) + \Sigma h_i(x_i)$

look variable $x_i$ alone, minimized at $x$

$\ell_i(x_i) = g(x) + h_i(x_i)$ thought of only in terms of $x_i$

hence $0 \in \partial \ell_i(x_i)$

$$0 \in \frac{\partial g}{\partial x_i} + \partial h_i(x_i)$$

$$-\frac{\partial g}{\partial x_i} \in \partial h(x_i)$$

hence $h_i(y_i) \geq h_i(x_i) - \frac{\partial g}{\partial x_i}(y_i - x_i)$