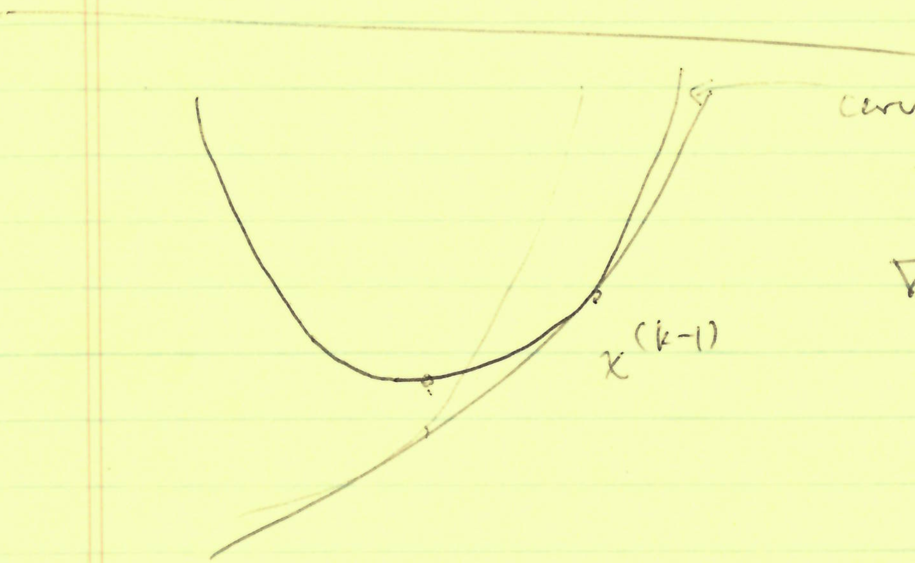
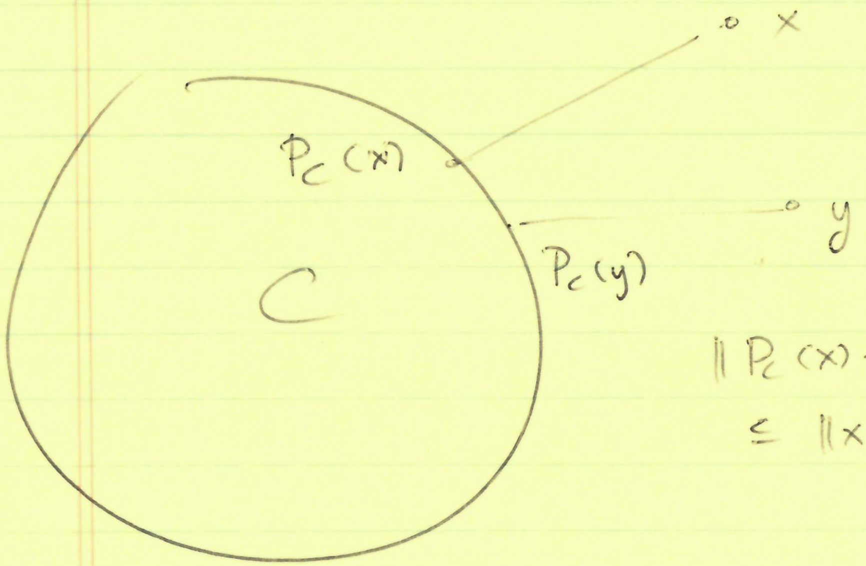
$M^{(k)}$ easily updated from $M^{(k-1)}$



curvature:

$\frac{1}{t} I$ (gradient)

$\nabla^2 f(x)$ (Newton)



$\|P_C(x) - P_C(y)\|_2 \le \|x - y\|_2$

Prox Newton: iteratively min.

$$b^T x + \frac{1}{2} x^T A x + \underline{\underline{h}}(x)$$

(2)

$$\|x^+ - x^\star\|_H = \|\operatorname{prox}_H(x - H^{-1}\nabla g(x)) - \operatorname{prox}_H(x^\star - H^{-1}\nabla g(x^\star))\|_H$$

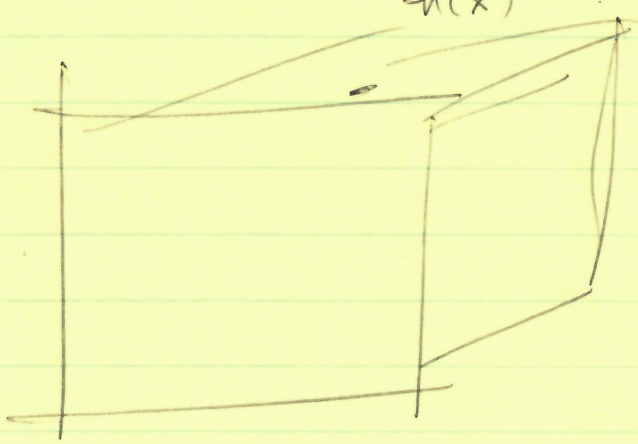
$$\tilde{g}_{k-1}(z) = \nabla g(x)^T(z-x) + \frac{1}{2t}\|z-x\|_2^2$$

---

unscaled prox

$$\operatorname{prox}_h(x - t\nabla g(x)) = x - t \cdot G_t(x)$$

$$\min_x \; g(x) + \underbrace{I_C(x)}_{h(x)}$$



$$z^{(k)}_{F_{k-1}} = x^{(k-1)}_{F_{k-1}} - t_k S^{(k-1)} \cdot \nabla_{F_{k-1}} g(x^{(k-1)})$$

$$z^{(k)}_{B_{k-1}} = x^{(k-1)}_{B_{k-1}}$$

$$x^{(k)} = P_{[l,u]}(z^{(k)})$$

3

glmnet

$$\underset{\beta}{\operatorname{argmin}} \quad g(\beta) + \lambda \|\beta\|_1$$

$$\iff \underset{\beta}{\operatorname{arg\,min}} \quad L(X\beta) + \lambda \|\beta\|_1$$

L = logistic regression loss

solves over $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_m \geq 0$

$$\text{KKT} \begin{cases} -X_{\cdot 1}^T (y - p(\beta)) = \lambda s_1 \\ - \quad \vdots \\ - \\ -X_p^T (y - p(\beta)) = \lambda s_p. \end{cases} \qquad p_i(\beta) = \frac{1}{1 + \exp(-(X\beta)_i)}$$

$$A = \{ j \in \{1, \ldots p\} : |X_j^T (y - p(\beta^{(0)}))| \geq \lambda \}$$

- proximal Newton only on vars in A
- upon convergence, recheck KKT over all variables
- if any vars fail their KKT, then add to A and repeat