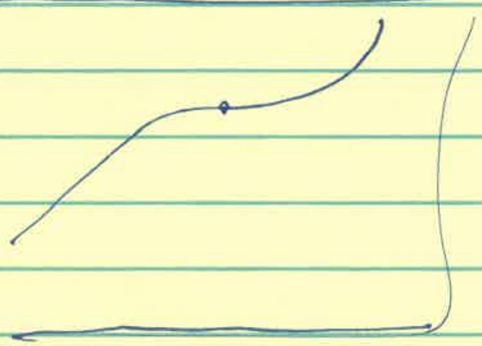
$$0 \ni \partial_x \left(f(x) + \sum u_i h_i(x) + \sum v_j l_j(x) \right).$$

$$\min_x L(x, u^*, v^*) = L(x^*, u^*, v^*)$$

by subgradient optimality

$$\Longleftrightarrow \quad 0 \in \partial_x \left(f(x^*) + \sum u_i^* h_i(x^*) + \sum v_j^* l_j(x^*) \right)$$

i.e. stationarity.



$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h_i(x) \le 0 \\ & l_j(x) = 0 \end{aligned}$$

$$\Longleftrightarrow \min \; f(x) + \sum 1\{h_i(x) = 0\} + \sum 1\{l_j(x) = 0\}$$

$$\begin{aligned} \min \quad & \tfrac{1}{2} x^T Q x + c^T x \,. \quad Q \succ 0 \\ \text{s.t.} \quad & Ax = 0. \end{aligned}$$

unconstrained

$Qx = -c$

$x = -Q^{-1} c.$

$$L(x,u) = \tfrac{1}{2} x^T Q x + c^T x + u^T A x.$$

KKT conditions:

- stat. $\nabla_x L(x,u) = 0.$
  $\boxed{Qx + c + A^T u = 0.}$
- comp. slack: $\emptyset$
- feasibility: $\boxed{Ax = 0.}$

KKT matrix

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = \begin{bmatrix} -c \\ 0 \end{bmatrix}$$

$$\min_{x} \quad -\sum \log(\alpha_i + x_i)$$

$$\text{s.t.} \quad x \geq 0 \quad 1^T x = 1.$$

$$L(x,u,v) = -\sum \log(\alpha_i + x_i) - \sum u_i x_i + v(1^T x - 1)$$

KKT:

- stat. $0 = \nabla_x L(x,u,v)$

  for $i = 1 \ldots n$: $0 = -\frac{1}{\alpha_i + x_i} - u_i + v$

- comp. slack: $u_i x_i = 0.$ $\quad i = 1, \ldots n$
- feasibility: $x \geq 0, \quad 1^T x = 1$

  $u \geq 0.$

lead to an algorithm for computing x^*.

$$\min_{\beta, \beta_0, \xi} \quad \frac{1}{2}\|\beta\|_2^2 + C\sum \xi_i$$

$$\text{s.t.} \quad \xi_i \geq 0 \qquad v_i$$

$$y_i(x_i^T\beta + \beta_0) \geq 1 - \xi_i \qquad w_i$$

$$L(\beta, \beta_0, \xi, v, w) = \frac{1}{2}\|\beta\|_2^2 + C\sum \xi_i - \sum v_i \xi_i + \sum w_i \left(1 - \xi_i - y_i(x_i^T\beta + \beta_0)\right)$$

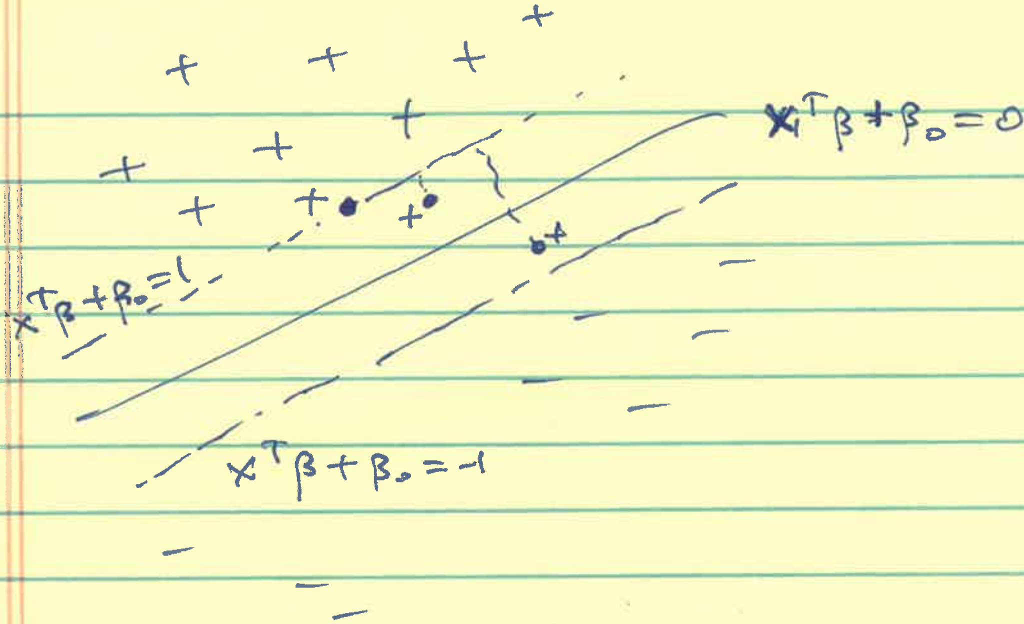
KKT:

- stat. $0 = \nabla_\beta L(\beta, \beta_0, \xi, v, w)$

  $0 = \beta - \sum w_i y_i x_i \Rightarrow \boxed{\beta = \sum w_i y_i x_i}$

$0 = \nabla_{\beta_0} L$ $\qquad$ $0 = \nabla_\xi L$

$\boxed{0 = \sum w_i y_i}$ $\qquad$ $\boxed{0 = C1 - v - w}$



$$L(x,\lambda) = f(x) + \lambda(h(x)-t)$$

(c) to (L) KKT: stat: x^* must solve $\min_x f(x)+\lambda(h(x)-t)$

$$\iff \min_x f(x)+\lambda h(x)$$

(L) to (c) KKT: stat: x^* must solve $\min_x f(x)+\lambda(h(x)-t)$

comp slack: $\lambda\cdot(h(x^*)-t)=0.$ note: satisfied by $t=h(x^*)$.

feasibility: $h(x^*)\le t.$

$\lambda \geq 0.$

---

eg $f(u) = \frac{1}{2}\|y-u\|_2^2$ so $f(X\beta)=\frac{1}{2}\|y-X\beta\|_2^2$

or logistic loss

or Poisson loss.

$$\partial_\beta \left(f(X\beta) + \lambda \|\beta\|_1 \right) \ni 0.$$

$$X^T f(X\beta) + \lambda s = 0. \quad \text{where } s \in \partial \|\beta\|_1.$$

$$s_i X_i = \sum \underbrace{s_i s_j c_j}_{\underline{\underline{a_j}}} (X_j \cdot s_j)$$

deduce $\sum a_j = 1$.