2/2 Recitation

- KKT conditions
- QPAL from HW 1

for convex:

$f(x) \to$ convex

$$\min_x \quad f(x)$$

$$\text{s.t.} \quad c(x) = 0 \quad : \lambda$$

$$g(x) \le 0 \quad : \mu$$

this problem is convex if f(x) is convex, and the constraints define a convex set. This means c(x) must be linear (Ax -b), and g(x) is convex

$$L(x, \lambda, \mu) = f(x) + \lambda^T c(x) + \mu^T g(x)$$

here is the Lagrangian

Stationarity: $\nabla_x L = \nabla_x f(x) + \left(\frac{\partial c}{\partial x}\right)^T \lambda + \left(\frac{\partial g}{\partial x}\right)^T \mu = 0$

primal feasibility: $c(x) = 0$

$g(x) \le 0$

dual feas.: $\mu \ge 0$

Compl. slackness: $\mu_i \cdot g_i(x) = 0 \quad \forall i$ ①

$\mu \circ g(x) = 0$ ☆ ②

$\mu^T g(x) = 0$ ③

You will see the complementary slackness written out in these 3 ways. I prefer the first and second ways to the third, though if 1 and 2 are true, then 3 is true as well.

$Q \in S^+$, $\mathrm{eig}(Q) \geq 0$, PSD

this is a convex QP if Q is positive semi definite (eig(Q)>=0)

does not change solution

we can omit any constant terms in the cost function since they will not change the minimizing argument

$$\min_x \quad \frac{1}{2} x^T Q x + q^T x + \cancel{r}$$

$$\text{s.t.} \quad Ax - b = 0 \quad : \lambda$$

$$Gx - h \leq 0 \quad : \mu$$

$$\underline{x} \leq x \leq \overline{x}$$

$$\Downarrow$$

$$x \leq \overline{x}$$

$$-x \leq -\underline{x}$$

$$\Downarrow$$

here is an example of how to turn bound constraints into the Gx<=h form that we need.

$$\underbrace{\begin{bmatrix} I \\ -I \end{bmatrix}}_{G} x \leq \underbrace{\begin{bmatrix} \overline{x} \\ -\underline{x} \end{bmatrix}}_{h}$$

Lagrangian

$$L(x, \lambda, \mu) = \frac{1}{2} x^T Q x + q^T x + \lambda^T (Ax - b) + \mu^T (Gx - h)$$

KKT conditions

$$\nabla_x L = Qx + q + A^T \lambda + G^T \mu = 0$$

$$Ax - b = 0$$

$$Gx - h \leq 0$$

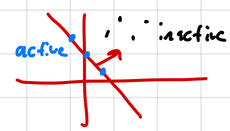
$$\mu \geq 0$$

$$\mu \odot (Gx - h) = 0$$

QP AL solver

$x \geq 0$

$x^p = 0$



Here is our Augmented Lagrangian. It is the original Lagrangian with added quadratic penalty terms for the equality and inequality constraints.

$$\text{AL:} \quad L_\rho(x, \lambda, \mu, \rho) = L(x, \lambda, \mu) + \frac{1}{2} \rho \overbrace{(Ax - b)}^{eq}{}^T (Ax - b) + \frac{1}{2} \overbrace{(Gx - h)}^{ineq}{}^T I_\rho (Gx - h)$$

$$I_\rho \in S^{\#\text{ of inequality}}$$

I_ρ: zeros(,)

for $i = 1 : \#\,ineq$

  if $(Gx - h)_i < 0$ and $\mu_i = 0$

    $I_\rho[i,:] = 0$

  else

    $I_\rho[i,:] = \rho$

I_\rho is a diagonal “mask matrix” or “selection matrix”, it is used to determine which of the inequality constraints is active and needs a penalty. If the constraint is not active, and the dual variable for that constraint is 0, then we can disreguard the constraint by making I_\rho(i,i) = 0.

putting L_ρ into a quadratic:

here is a demo of how I put the AL into a quadratic form that makes taking derivatives easy

$$L_\rho = \underbrace{\tfrac{1}{2}x^TQx + q^Tx + \lambda^T(Ax-b) + \mu^T(Gx-h)}_{1} + \underbrace{\tfrac{1}{2}\rho(Ax-b)^T(Ax-b)}_{2} + \underbrace{\tfrac{1}{2}(Gx-h)^T I_\rho (Gx-h)}_{3}$$

1 $\frac{1}{2}x^TQx + (q + A^T\lambda + G^T\mu)^T x + \dots$

Since we don't care about constant terms, we can ignore them (hence the …'s)

2 $\frac{1}{2}x^T(\rho A^TA)x + (-\rho A^Tb)^T x + \dots$

3 $\frac{1}{2}(Gx-h)^T I_\rho (Gx-h)$

$\frac{1}{2}(x^TG^T - h^T)(I_\rho Gx - I_\rho h)$

$\frac{1}{2}(x^TG^TI_\rho Gx - h^TI_\rho Gx - x^TG^TI_\rho h + \dots)$

$\frac{1}{2}x^T(G^TI_\rho G)x + (-G^TI_\rho h)^T x + \dots$

If any of these reduction steps confuses you, see the 1/27 recitation.

$$L_\rho(x,\lambda,\mu,\rho) = \frac{1}{2}x^T(Q + \rho A^TA + G^TI_\rho G)x + (q + A^T\lambda + G^T\mu - \rho A^Tb - G^TI_\rho h)^T x + \dots$$

$$\nabla_x L_\rho = (Q + \rho A^TA + G^TI_\rho G)x + q + A^T\lambda + G^T\mu - \rho A^Tb - G^TI_\rho h$$

$$= Qx + q + A^T(\lambda + \rho(Ax-b)) + G^T(\mu + I_\rho(Gx-h))$$

$$\nabla_x^2 L_\rho = Q + \rho A^TA + G^TI_\rho G$$

Gradient and Hessian of AL

QPAL alg

Here is the QPAL alg you will implement in HW1/Q3

init: $x = 0$, $\lambda = 0$, $\mu = 0$, $\rho = 1$, $\phi = 10$

for $i = 1:$ max_AL_iter

① solve $x = \min_x L_\rho(x, \lambda, \mu, \rho)$

hold these constant

Use Newton's method, $\alpha = 1$ is ok

We use Newton's method to solve this unconstrained minimization of the AL wrt x. NOTE: you can treat I_rho as constant when taking derivatives of the AL, but make sure you update I_rho every time you update x.

② update dual's

$$\lambda = \lambda + \rho(Ax - b)$$

$$\mu = \max(0, \mu + \rho(Gx - h))$$

③ update ρ

$$\rho = \rho \cdot \phi$$

④ Check convergence

$\|Ax - b\|_\infty \Rightarrow \underbrace{\text{norm}(\ , \text{Inf})}_{\text{Julia}}$

$\max(0, \text{maximum}(Gx - h))$

You can compute constraint violations for equality with norm(Ax-b,Inf), and inequality with max(0,maximum(Gx-h)) in Julia. The max of these should be below some tolerance specified in the function. Alternatively, you can use the full KKT conditions as termination criteria.

or check KKT