- Review eigenvectors / eigenvalues $\Longleftarrow$
- optimization : quadratic functions $\bar{w}$ linear consts.

OPTIMIZATION: (UNCONSTRAINED)

$$\min_{x \in \mathbb{R}^n} f(x)$$

objective function:

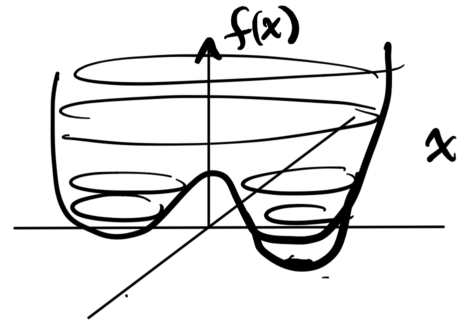
$$f : \mathbb{R}^n \longrightarrow \mathbb{R}$$

↳ encodes what we want to minimize.

f : differentiable

OPTIMALITY COND : $\dfrac{\partial f}{\partial x} = 0$

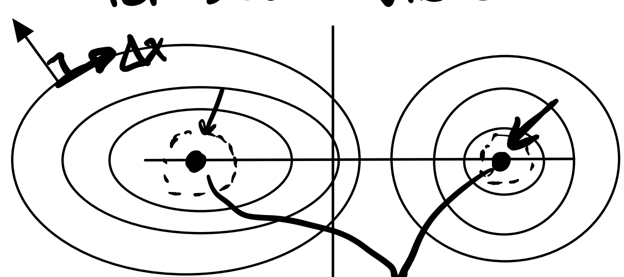
local optimality



Gradient points directly up hill...

TOP DOWN VIEW:

$\dfrac{\partial f}{\partial x} \perp$ to level sets



critical pts where $\dfrac{\partial f}{\partial x} = 0$

$$\Delta f = \frac{\partial f}{\partial x} \Delta x = 0$$

↳ perturbation along level set

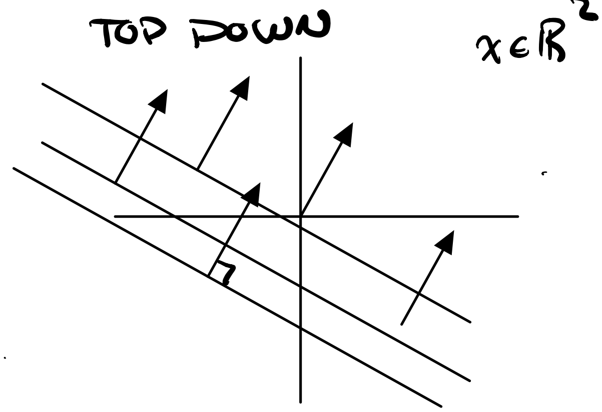
↙

Simple objective functions:

Linear obj.

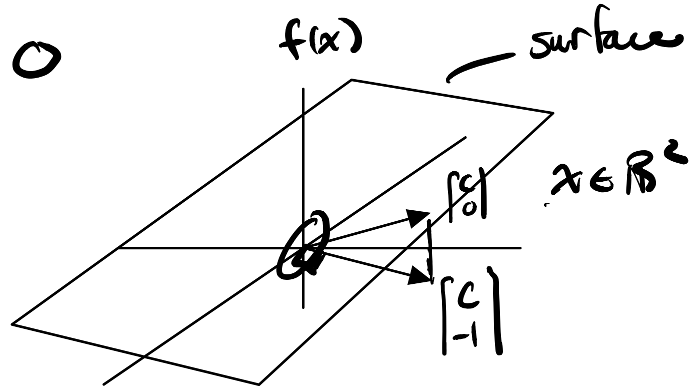
$$f(x) = c^T x$$

$$\frac{\partial f}{\partial x} = c^T$$



$$\begin{bmatrix} c^T & -1 \end{bmatrix} \begin{bmatrix} x \\ f \end{bmatrix} = 0$$

$$\min_{x \in \mathbb{R}^n} \; c^T x$$



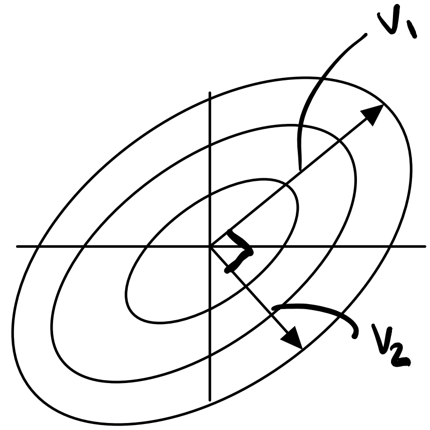
Quadratic Objective (canonical convex objective)

$$f(x) = \frac{1}{2} \underline{x^T Q x} + \underline{c^T x}$$

No linear term:

$$f(x) = \frac{1}{2} x^T Q x$$

ellipsoid shapes are determined by eigenvectors & eigen values of Q



for $f(x) = x^T Q x$ always assume Q is symmetric.

(if not: $x^T\left(\underbrace{\tfrac{1}{2}(Q+Q^T)}_{\text{sym}} + \underbrace{\tfrac{1}{2}(Q-Q^T)}_{\text{skew sym}}\right)x$)

replace Q w/ $\underbrace{\tfrac{1}{2}(Q+Q^T)}_{\text{sym component of } Q}$

$\frac{1}{2}x^TQx - \frac{1}{2}x^TQ^Tx = 0$

EIGENVALUE PROB: REVIEW

$Q\underline{v} = \underline{\lambda}\,\underline{v}$ $\quad Q \in \mathbb{R}^{n\times n}$ $\quad \underbrace{v \in \mathbb{R}^n}_{\text{eigenvector}}$, $\underbrace{\lambda \in \mathbb{R} \text{ or } \mathbb{C}}_{\text{eigenvalue}}$

bilinear

first solve for λ

$(\underbrace{\lambda I - Q}_{\text{have a nontrivial nullspace}})v = 0 \Rightarrow \det(\lambda I - Q) = 0$

$\chi_Q(s) = \det(sI - Q) = s^n + \underline{\alpha_{n-1}s^{n-1}} + \cdots + \alpha_1 s + \alpha_0$

→ roots of $\chi_Q(s)$ are eigenvalues

nth order → always has n roots

fund. thm of algebra

$\{\lambda_1, \dots, \lambda_n\} = \text{spec}(Q)$

for ea. $\lambda_i \Longrightarrow$ solve for v_i (eigenvector) $(\lambda_i I - Q)v_i = 0$

if we can find n linearly ind. eigenvectors then Q is diagonalizable

basis of eigenvectors

$$[Qv_1 \cdots Qv_n] = [\lambda_1 v_1 \cdots \lambda_n v_n]$$

$$Q \underbrace{[v_1 \cdots v_n]}_{V} = \underbrace{[v_1 \cdots v_n]}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}}_{D}$$

$$QV = VD$$

$$Q = VDV^{-1}$$

} $\to$ diagonalization of Q

Q is related to D by a similarity transform.

can think of Q in coords w.r.t. V as just stretching along the coord axes

Why do we care?

$$Q^k = VDV^{-1} \times VDV^{-1} \times \cdots \times VDV^{-1} = VD^kV^{-1}$$

$$= [V] \begin{bmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{bmatrix} V^{-1}$$

if $\lambda \in \text{spec}(Q) \implies \lambda^k \in \text{spec}(Q^k)$

for general analytic functions $f(\cdot)$

if $\lambda \in \text{spec}(Q) \implies f(\lambda) \in \text{spec}(f(Q))$

eigenvectors of Q^k, $f(Q)$, etc same as Q

$Q^{1/2} = M$ s.t. $MM = Q$

$$M = \underline{V} \begin{bmatrix} \lambda_1^{1/2} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{1/2} \end{bmatrix} \underline{V}^{-1} \qquad MM = VD^{1/2}V^{-1}VD^{1/2}V^{-1}$$
$$= VDV^{-1} = Q$$

Symmetric Matrices

$Q = Q^T$

- $\operatorname{spec}(Q) \in \mathbb{R}$ real eigenvalues
- eigenvectors are orthogonal to ea. other

v_i, v_j are eigenvectors of Q

$j \neq i \quad v_i^T v_j = 0$

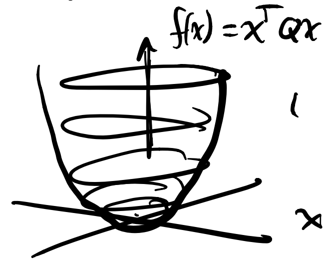
$$Q = RDR^T$$

$R = [v_1 \cdots v_n]$ $\quad \underline{R^{-1} = R^T}$ $\quad R$: unitary ...
orthonormal ...
rotation ...

$$\underline{R}^T R = \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} \begin{bmatrix} v_1 \cdots v_n \end{bmatrix} = \begin{bmatrix} \overset{1}{v_1^T v_1} & \cdots & v_1^T v_n \\ v_n^T v_1 & \cdots & \underset{1}{v_n^T v_n} \end{bmatrix} = I$$

off diagonal elements are 0

$R^T = R^{-1} \rightarrow$ orthonormal eigenvectors are perpendicular sign of eigenvalues of $Q = Q^T \ldots$

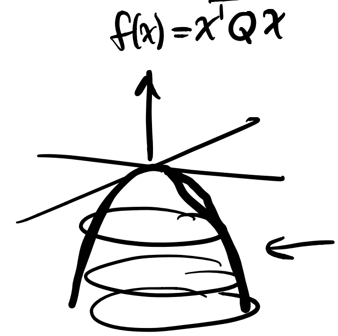


<u>Defn</u> Positive definite $Q \succ 0$

$Q = Q^T \quad x^T Q x > 0 \quad \forall x \in \mathbb{R}^n$

(positive semi definite) $Q \succeq 0$

$x^T Q x \geq 0 \quad \forall x \in \mathbb{R}^n$



<u>Defn</u> negative definite $Q \prec 0$

$Q = Q^T \quad x^T Q x < 0 \quad \forall x \in \mathbb{R}^n$

negative semi definite $Q \preceq 0$

$x^T Q x \leq 0 \quad \forall x \in \mathbb{R}^n$

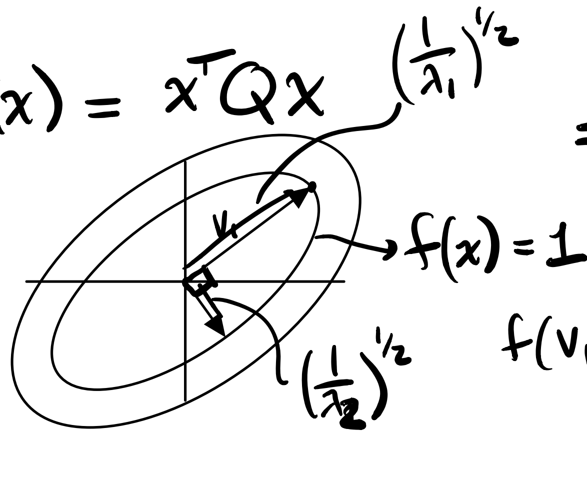
$$Q \succ 0 \iff \operatorname{spec}(Q) > 0$$

$$f(v_i) = v_i^T Q v_i = \underline{v_i}^T \begin{bmatrix} v_1 \cdots v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} \underline{v_i}$$

$$\begin{bmatrix} 0 & |v_i| & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} 0 \\ |v_i| \\ 0 \end{bmatrix}$$

$$f(x) = x^T Q x$$

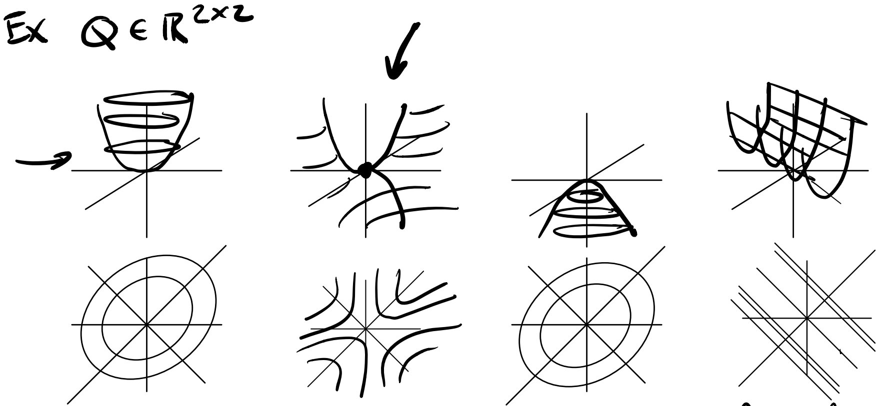
$$= \underline{|v_i|^2} \, \underline{\lambda_i} \rightarrow$$



$$f(v_1) = 1 = |v_1|^2 \lambda_1$$

$$|v_1| = \left(\frac{1}{\lambda_1}\right)^{1/2}$$

Ex $Q \in \mathbb{R}^{2\times 2}$

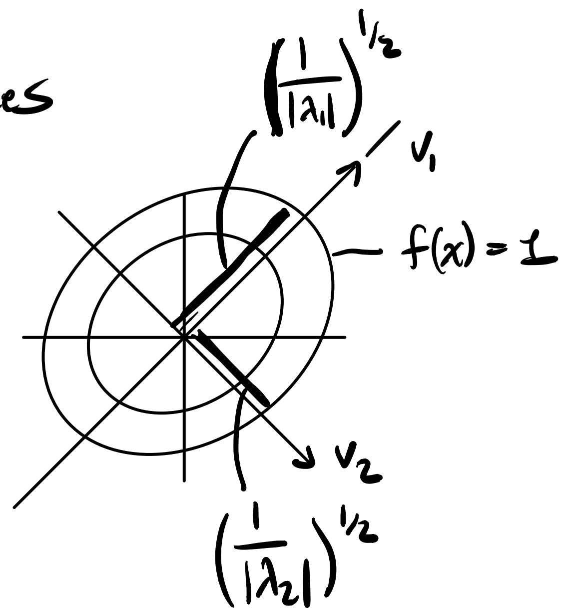


$\lambda_1, \lambda_2 > 0$	$\lambda_1 > 0,\ \lambda_2 < 0$	$\lambda_1, \lambda_2 < 0$	$\lambda_1 > 0,\ \lambda_2 = 0$
positive definite	indefinite	negative definite	positive semi definite

ML algorithms avoiding saddles

$$f(x) = x^T Q x$$

adding a linear term...



$$f(x) = x^T Q x + c^T x$$

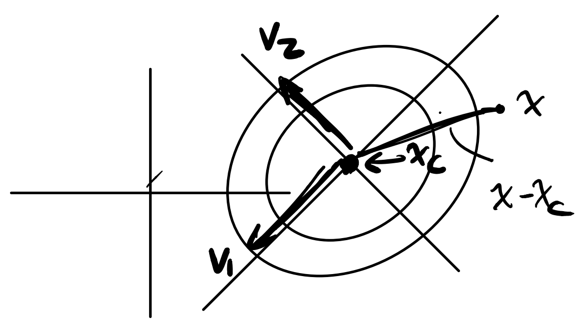
computing new center point...

$$\frac{\partial f}{\partial x} = 0 : \quad 2x^T Q + c^T = 0$$

$$x^T = -\tfrac{1}{2} c^T Q^{-1}$$

$$x^T = -\tfrac{1}{2} c^T R D^{-1} R^T =$$

$$= -\tfrac{1}{2} \underbrace{\begin{bmatrix} \frac{c^T v_1}{\lambda_1} \cdots \frac{c^T v_n}{\lambda_n} \end{bmatrix}} \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}$$



coords of center pt. wrt eigenvector basis

$$f(x) = x^TQx + c^Tx = (x - x_c)^TQ(x - x_c) + \text{const}$$

$x^TQx = (x-0)^TQ(x-0)$

$$= x^TQx - 2x_c^TQx + x_c^TQx_c + \text{const}$$

$\Rightarrow \text{const} = -x_c^TQx_c$

$c^Tx = -2x_c^TQx$

$c^T = -2x_c^TQ \Rightarrow x_c^T = -\frac{1}{2}c^TQ^{-1}$

Adding constraints....

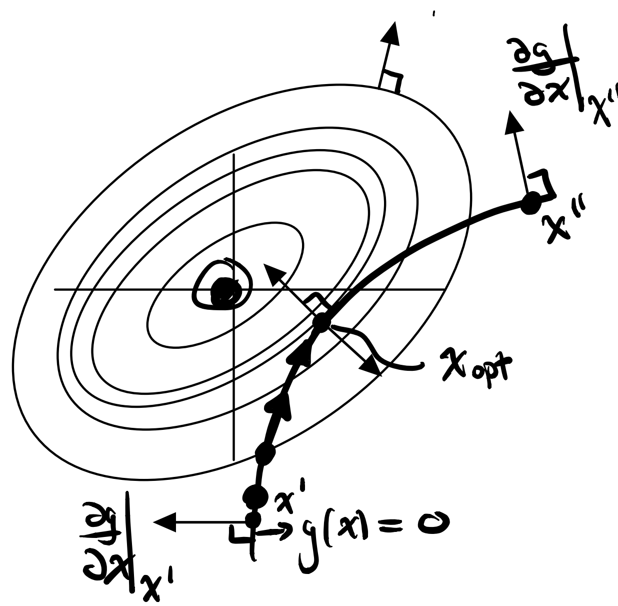
Equality constraints...

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } g(x) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$f: \mathbb{R}^n \to \mathbb{R}$ —

$g: \mathbb{R}^n \to \mathbb{R}^m$

$\frac{\partial g}{\partial x} \in \mathbb{R}^{m \times n}$



$\frac{\partial g}{\partial x} \in \mathbb{R}^{1 \times 2}$

what optimality condition tells us to stop at x_{opt}?

Option 1: constrain ourselves to $g(x)$
and optimize f on that set

$\Rightarrow$ converting a constrained optimization
into an unconstrained opt.

Option 2: optimality condition in terms
$\frac{\partial f}{\partial x}$ and $\frac{\partial g}{\partial x}$

Method of
Lagrange
multipliers

unconstrained: $\frac{\partial f}{\partial x} = 0$

key insight: " $\frac{\partial f}{\partial x}$ and $\frac{\partial g}{\partial x}$ point in the same subspace "

$$\boxed{\frac{\partial f}{\partial x} = v^T \frac{\partial g}{\partial x}}$$

$\frac{\partial f}{\partial x}$ is a lin comb of the rows of $\frac{\partial g}{\partial x}$

replaces $\frac{\partial f}{\partial x} = 0$

Examples Linear constraints $Ax = b$

$A \in \mathbb{R}^{m \times n}$ $b \in \mathbb{R}^m$

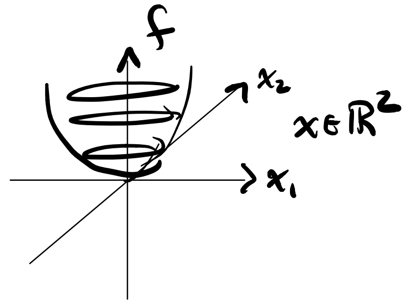
$$\min_x f(x)$$
$$\text{s.t. } g(x) = 0$$

$$g(x) = \underline{Ax - b} = 0$$

m constraints.

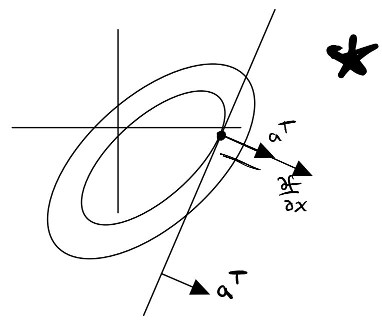
$$\frac{\partial g}{\partial x} = A \qquad g(x) = a^T x - b$$

$$\frac{\partial g}{\partial x} = a^T$$



Ex.

$$\min_{x \in \mathbb{R}^2} \frac{1}{2} x^T Q x + c^T x$$
$$\text{s.t. } a^T x = b$$



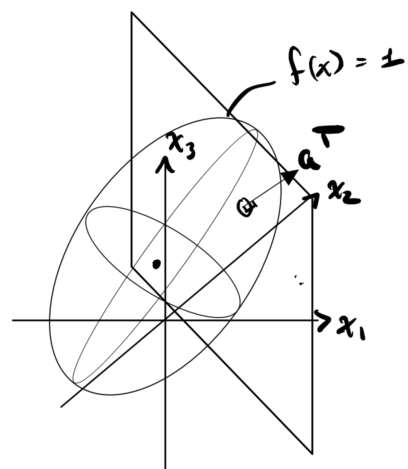
$$\frac{\partial f}{\partial x} = x^T Q + c^T$$

v : allows us to account for steepness of f

$$\frac{\partial f}{\partial x} = v a^T \quad v \in \mathbb{R}$$

solve together
$$\left[\begin{array}{l} x^T Q + c^T = v a^T \to \text{at opt point} \\ a^T x = b \to \text{on the line} \end{array}\right]$$

$$x^T = (v a^T - c^T) Q^{-1}$$

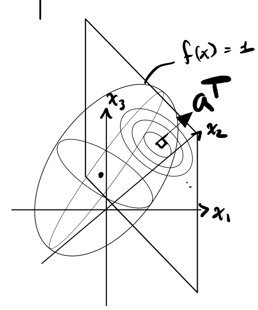


Ex.
$$\min_{x \in \mathbb{R}^3} \frac{1}{2} x^T Q x + c^T x$$
$$\text{s.t. } \underline{a^T x = b}$$

$$\frac{\partial f}{\partial x} = x^T Q + c^T = v a^T \quad v \in \mathbb{R}$$

$$a^T x = b$$

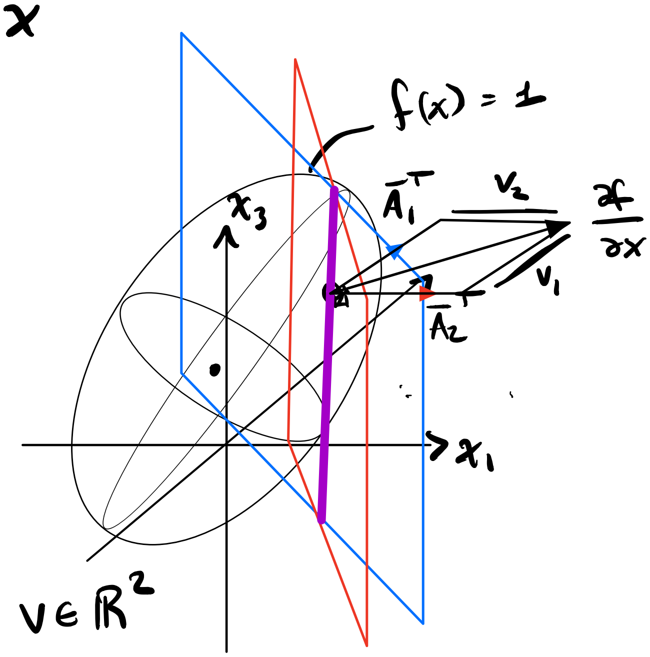
solve for x and v



Ex. $\min_{x \in \mathbb{R}^3} \frac{1}{2} x^T Q x + c^T x$

s.t. $Ax = b$

$$A = \begin{bmatrix} -\bar{A}_1^T- \\ -\bar{A}_2^T- \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$



$$\frac{\partial f}{\partial x} = v^T \frac{\partial g}{\partial x}$$

$$\downarrow$$

$$\frac{\partial f}{\partial x} = [v_1 \; v_2] \begin{bmatrix} -\bar{A}_1^T- \\ -\bar{A}_2^T- \end{bmatrix} \qquad v \in \mathbb{R}^2$$

$$= v_1 \bar{A}_1^T + v_2 \bar{A}_2^T$$

$$\left.\begin{aligned} x^T Q + c^T &= v^T A \\ Ax &= b \end{aligned}\right\} \longrightarrow \text{solve for } x \;\&\; v$$

Summary

$$\frac{\partial f}{\partial x} = v^T \frac{\partial g}{\partial x} \quad \Leftarrow \text{ stationarity (gradient is 0)}$$

$$g(x) = 0 \quad \Leftarrow \text{ feasibility (within the feasible set)}$$

solve for $\underline{x}$ and $\underline{v}$

Solving for x & v

$$\min_x \ \frac{1}{2}x^TQx + c^Tx$$
$$\text{s.t.} \quad Ax = b$$

$$\frac{\partial f}{\partial x} = v^T\frac{\partial g}{\partial x} \Rightarrow \boxed{x^TQ + c^T = -v^TA}$$
$$g(x) = 0 \Rightarrow \boxed{Ax = b}$$

→ sys of lin eqns.

$$\underbrace{\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix}}_{\text{want to invert.}} \underbrace{\begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}}_{\text{solve}}$$

the initial direction we assign to v doesn't matter

a simpler problem: $B \in \mathbb{R}^{m \times n}$ $m < n$

Block matrix inversion

$$\begin{bmatrix} I & B^T \\ B & 0 \end{bmatrix}^{-1} \begin{bmatrix} I & B \\ B & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

$$\begin{bmatrix} * & * \\ * & * \end{bmatrix} \begin{bmatrix} I & B \\ B & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

B: $\Rightarrow$ find basis for nullspace $\Rightarrow$ cols of M

$$\underline{BM = 0} \quad \& \quad \mathcal{R}(M) = \mathcal{N}(B)$$

$$[B^T\ M]^{-1} = \begin{bmatrix}(BB^T)^{-1}B \\ (M^TM)^{-1}M^T\end{bmatrix} \qquad \begin{bmatrix}(BB^T)^{-1}B \\ (M^TM)^{-1}M^T\end{bmatrix}[B^T\ M] = \begin{bmatrix} I & 0 \\ 0 & I\end{bmatrix}$$

$$[B^T\ M]\begin{bmatrix}(BB^T)^{-1}B \\ (M^TM)^{-1}M^T\end{bmatrix}$$

$$\underbrace{B^T(BB^T)^{-1}B}_{\text{proj}_{B^T}} + \underbrace{M(M^TM)^{-1}M^T}_{\text{proj}_M} = I$$

$$\begin{bmatrix} I & B^T \\ B & 0\end{bmatrix}\begin{bmatrix} * & B^T(BB^T)^{-1} \\ * & * \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I\end{bmatrix}$$

$$\begin{bmatrix} I & B^T \\ B & 0\end{bmatrix}\begin{bmatrix} * & B^T(BB^T)^{-1} \\ * & -(BB^T)^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I\end{bmatrix}$$

$$\begin{bmatrix} I & B^T \\ B & 0\end{bmatrix}\begin{bmatrix} M(M^TM)^{-1}M^T & B^T(BB^T)^{-1} \\ * & -(BB^T)^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I\end{bmatrix}$$

$$\begin{bmatrix} I & B^T \\ B & 0\end{bmatrix}\begin{bmatrix} M(M^TM)^{-1}M^T & B^T(BB^T)^{-1} \\ (BB^T)^{-1}B & -(BB^T)^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I\end{bmatrix}$$

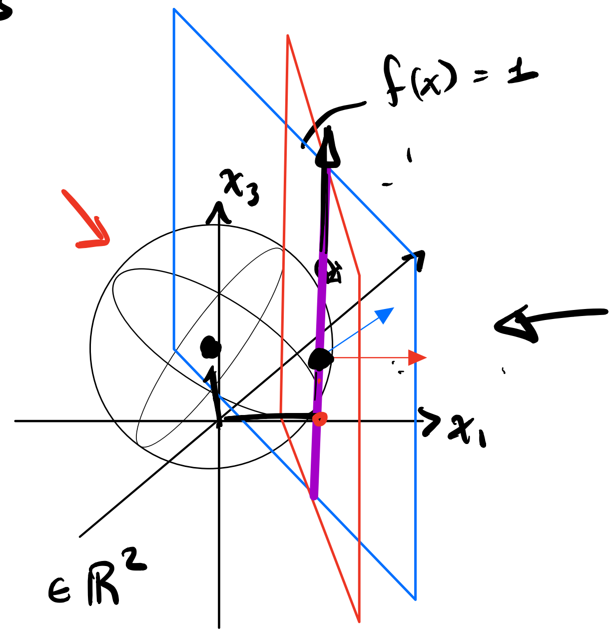
$$\Rightarrow \begin{bmatrix} I & B^T \\ B & 0 \end{bmatrix}^{-1} = \begin{bmatrix} M(M^TM)^{-1}M^T & B^T(BB^T)^{-1} \\ (BB^T)^{-1}B & -(BB^T)^{-1} \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} I & B^T \\ B & 0 \end{bmatrix}\begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}}_{\text{optimality conds for}}$$

$$\min_x \ \tfrac{1}{2}x^Tx + c^Tx$$

$$\underbrace{\text{s.t. } Bx = b}$$

level sets of $\frac{1}{2}x^Tx + c^Tx$ are spheres



Solution:

$$\begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} I & B^T \\ B & 0 \end{bmatrix}^{-1}\begin{bmatrix} -c \\ b \end{bmatrix}$$

$$= \begin{bmatrix} M(M^TM)^{-1}M^T & B^T(BB^T)^{-1} \\ (BB^T)^{-1}B & -(BB^T)^{-1} \end{bmatrix}\begin{bmatrix} -c \\ b \end{bmatrix}$$

$$\Rightarrow \quad \begin{aligned} x &= -M(M^TM)^{-1}M^Tc + B^T(BB^T)^{-1}b \\ v &= (BB^T)^{-1}Bc - (BB^T)^{-1}b \end{aligned}$$

$$x = -\underbrace{M(M^TM)^{-1}M^Tc}_{\text{proj}_M c} + \underbrace{B^T(BB^T)^{-1}b}_{\text{minimum norm } x \text{ s.t. } Bx = b}$$

$x = Q^{-1/2} x'$: coord transform → turns level set ellipsoids into spheres

$$\min_x \ \tfrac{1}{2} x^T Q x + c^T x \quad \text{s.t.} \quad Ax = b \implies \min_{x'} \ \tfrac{1}{2} x'^T x' + c^T Q^{-1/2} x' \quad \text{s.t.} \quad \underbrace{AQ^{-1/2}}_{B} x' = b$$

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix} = \begin{bmatrix} Q^{1/2} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & Q^{-1/2}A^T \\ AQ^{-1/2} & 0 \end{bmatrix} \begin{bmatrix} Q^{1/2} & 0 \\ 0 & I \end{bmatrix}$$

$$\boxed{B = AQ^{-1/2}} \qquad \boxed{M = Q^{1/2} N} \qquad R(N) = N(A)$$

$$(BM = AQ^{-1/2} Q^{1/2} N = 0)$$

proj.

from above

$$\begin{bmatrix} I & B^T \\ B & 0 \end{bmatrix}^{-1} = \begin{bmatrix} M(M^TM)^{-1}M^T & B^T(BB^T)^{-1} \\ (BB^T)^{-1}B & -(BB^T)^{-1} \end{bmatrix}$$

$$\begin{bmatrix} I & Q^{-1/2}A^T \\ AQ^{-1/2} & 0 \end{bmatrix}^{-1} = \begin{bmatrix} Q^{1/2}N(N^TQN)^{-1}N^TQ^{1/2} & Q^{-1/2}A^T(AQ^{-1}A^T)^{-1} \\ (AQ^{-1}A^T)^{-1}AQ^{-1/2} & -(AQ^{-1}A^T)^{-1} \end{bmatrix}$$

$$\begin{bmatrix} Q & A^T \\ A & 0 \end{bmatrix}^{-1} = \begin{bmatrix} N(N^TQN)^{-1}N^T & Q^{-1}A^T(AQ^{-1}A^T)^{-1} \\ (AQ^{-1}A^T)^{-1}AQ^{-1} & -(AQ^{-1}A^T)^{-1} \end{bmatrix}$$

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

$Ax = b$

$$x = -\underline{N(N^T Q N)^{-1} N^T} c + \underline{Q^{-1} A^T (A Q^{-1} A^T)^{-1}} b$$

$$v = -(A Q^{-1} A^T)^{-1} A Q^{-1} c - (A Q^{-1} A^T)^{-1} b$$

$A^T (A A^T)^{-1}$ $\quad$ $A (A^T A)^{-1}$ $\Leftarrow$

$A A^T A^T$ $\quad$ $(A Q^{-1} A^T) A$

$A^T (A^T A)^{-1}$