

CHAPTER 1: BASICS OF NON-LINEAR

SOLUTIONS

GLOBAL
LOCAL

$$f(x^*) \leq f(x) \quad \forall x$$

$$f(x^*) \leq f(x) \quad \forall x \in S = \|x - x^*\| \leq \epsilon$$

* STRICT

WITHOUT EQUALITY

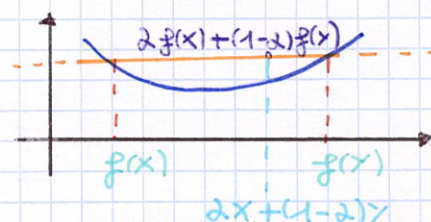
CONVEXITY

SET
FUNCTION

$$\alpha x + (1-\alpha)y \in S \quad \forall x, y \in S, \alpha \in [0, 1]$$

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$$

$$\forall x, y \in S, \alpha \in [0, 1]$$



OR THE SEGMENT LAYS ABOVE THE FUNCTION

THE CONVEX OPTIMIZATION THEOREM

THM LET x^* BE A LOCAL MINIMIZER OF $\min_{x \in S} f(x)$: S IS CONVEX. THEN, x^* IS A GLOBAL MINIMIZER. IF $f(x)$ IS STRICTLY CONVEX, THEN IT IS UNIQUE

PROOF ASSUME $\exists y : f(y) < f(x^*)$ THEN

$$f(\alpha x^* + (1-\alpha)y) \leq \alpha f(x^*) + (1-\alpha)f(y)$$

$$< \alpha f(y) + (1-\alpha)f(x^*)$$

x^* NOT LOCAL MINIMIZER!

MATRIX SIGN

LET $A \in \mathbb{R}^{m \times m}$ BE A MATRIX WITH m ORTHONORMAL VECTORS $[v_1, \dots, v_m] = V$

ORTHONORMAL EIGENVALUES $v_i v_i^T = 1, v_i^T v_j = 0 \quad \forall i \neq j$
 $AV = [\lambda_1 v_1, \lambda_2 v_2, \dots, \lambda_m v_m]$

A IS SDP $\lambda_i \geq 0 \quad \forall i$ A IS HERMITIAN!
 (DP) $\lambda_i > 0 \quad \forall i$
 (NDP) $\lambda_i \leq 0 \quad \forall i$

SYLVESTER' RULE A IS DP IF ALL ITS m PRINCIPAL MINORS ARE POSITIVE

// ND IFF $-A$ IS POSITIVE DEF!

$$\det [a_{ij}], \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \dots$$

!! NO SDP OR SDN !!
 $\left. \begin{array}{l} +, +, +, + \text{ DP} \\ -, +, -, + \text{ ND} \\ \text{WRONG SIGN} \text{ IND} \end{array} \right\} > 0. \text{ IF } = 0 \text{ ?}$

A IS SDP ITS $2^m - 1$ MINORS ARE ≥ 0

DIRECTION.
DERIVATE

GIVEN $f: \mathbb{R}^m \rightarrow \mathbb{R}$ AND $x, d \in \mathbb{R}^m$

$$\nabla_d f(x) = f'(x; d) = \lim_{t \rightarrow 0} \frac{f(x+td) - f(x)}{t}$$

NOTE $\nabla_d f(x) = d^T \nabla f(x) = \|d\| \cdot \|\nabla f(x)\| \cdot \cos(\theta)$
 $\nabla f(x)$ IS A COLUMN VECTOR! $\cos \theta = \frac{a \cdot b}{\|a\| \|b\|}$ ^{ANGLE!}

DERIVATE
RECAP $(fg)' = f'g + fg'$ $(f/g)' = (f'g - fg')/g^2$

SPECIFIC $\frac{d}{dt} g(x+td) = g'_x(x+td) \cdot d$

JACOBIAN $f: \mathbb{R}^m \rightarrow \mathbb{R}^m$ THEN

$$\nabla^T f(x) = J = \begin{bmatrix} \frac{\partial f_1}{\partial x} \\ \vdots \end{bmatrix}$$

DERIVATES
AND CONVEX

SDP HESSIAN $x^T \nabla^2 f(x) x \geq 0 \quad \forall x$

POSITIVE EIG ≥ 0 ! NOT AN IFF!

$$H = \begin{bmatrix} \partial^2 f / \partial x_i \partial x_j \end{bmatrix}$$

THE GENER.
OPTIMIZATION
ALGORITHM

4. ALGORITHM 1. GUESS AN x_0

2. FOR $k=0, 1, \dots$

IF x_k OPT THEN STOP

DETERMINE DIRECTION p_k
STEPSIZE α_k

SO THAT $x_{k+1} = x_k + \alpha_k p_k$

DIRECTION

$C^T(x_k + \epsilon p_k) \leq C^T(x_k)$
FEASIBILITY $x_k + \alpha p_k \in S$

CONVERG.
RATE

THE RATE

SEQUENCE $\{x_k\}$ CONVERGE
TO x^* AT RATE ν IFF

$$\lim_{k \rightarrow \infty} \frac{\|e_{k+1}\|}{\|e_k\|^\nu} = C \in \mathbb{R}$$

WITH $e_k = x_k - x^*$

$\nu > 1$ SUPERLINEAR CONVERGENCE

$\nu = 1$ LINEAR

$\nu = 2$ QUADRATIC

TAYLOR SERIES

REAL FN $\sum_{m=0}^K \frac{f^{(m)}(x_0)}{m!}$ // SIGMA NOTATION

MULTIDIM $f(x_0+d) \approx f(x_0) + d^T \nabla f(x_0) + \frac{1}{2} d^T \nabla^2 f(x_0) d$
 $f(x+d) \approx \sum_{m=0}^K \frac{\nabla^{(m)} f(x_0)}{m!} \quad !! J(x) = \nabla f(x) !!$

LIPSCHITZ $f(x)$ IS LIPSCHITZ IN $x \in D$ WITH CONSTANT $L > 0$ IFF

$$|f(x_1) - f(x_2)| \leq L \|x_1 - x_2\| \quad \forall x_1, x_2 \in D$$

THM IF $\nabla f(x)$ IS LIPSCHITZ OF L , THEN
 $E(x+d) \leq \frac{L}{2} \|d\|^2$

PROOF DEF $E(x+d) := \|f(x+d) - T(x+d)\|$
 $= \|f(x+d) - f(x) + \nabla f(x) \cdot d\|$
 $= \left\| \int_0^1 \nabla f(x+td) dt - \nabla f(x) \right\| \|d\|$
DOES NOT DEP ON t INSERT THIS!
 $= \left\| \int_0^1 (\nabla f(x+td) - \nabla f(x)) dt \right\| \|d\|$
 $\leq \int_0^1 \|\nabla f(x+td) - \nabla f(x)\| \cdot \|d\| dt$
 $\leq \int_0^1 L \cdot \|x+td - x\| \cdot \|d\| dt$
 $= L \|d\|^2 \int_0^1 t dt = \frac{L}{2} \|d\|^2$

NEWTON

USES THE FIRST TWO TERMS OF TAYLOR TO APPROXIMATE A SOLUTION FOR F TANGENT LINE

A. INITIALIZE $x_0 \in \mathbb{R}^n, k \rightarrow 0$

B. LOOP FOR $k=0, 1, \dots$ MATRIX

$$x_{k+1} = x_k - \nabla^2 f(x_k)^{-1} \cdot \nabla f(x_k)$$

$k++$ $-1 \rightarrow T \rightarrow J_F^{-1}$

THM ASSUME 1. $f \in C^2$ 2. $\nabla^2 f$ LIPSCHITZ
 IF x^* IS A LOCAL MINIMUM, AND $\|x_0 - x^*\| \leq \epsilon$
 THEN NEWTON CONVERGES QUADRATICALLY

CAN FAIL! FOR INSTANCE, $f(x) = \tan(x)$

DESCENT DIRECTION

DEF $d \in \mathbb{R}^n$: $\nabla f(x) \cdot d < 0$
 $d^T \nabla f(x) < 0$

SINGLE FUNCTION ZERO

$$X_{k+1} = X_k - \frac{f(X_k)}{f'(X_k)} \text{ UNTIL } \|f(X_k)\| \leq \epsilon$$

SYSTEM OF EQUATIONS ZERO

$$X_{k+1} = X_k - J_F^{-1}(X_k) \cdot F(X_k)$$

NO ONES INVERTS...

$$J(X_k) \cdot (X_{k+1} - X_k) = -F(X_k)$$

NEWTON FOR MINIMIZATION

FIND A ZERO OF ∇f INSTEAD OF $x(y) = f(x)$

SINGLE FUNCTION

$$X_{k+1} = X_k - \frac{f'(X_k)}{f''(X_k)}$$

SYSTEMS OF EQUATIONS

$$X_{k+1} = X_k - (\nabla^2 f(X_k))^{-1} \nabla f(X_k) \quad O(m^3) \text{ FOR } dk$$

AGAIN...

$$\begin{cases} X_{k+1} = X_k + P_k \\ P_k \Rightarrow \nabla^2 f(X_k) \cdot P_k = -\nabla f(X_k) \end{cases} \quad \text{STEP OR DIRECTION}$$

CONVERGENCE THM

LET f BE DEFINED OVER THE CONVEX SET S . ALSO, LET $\nabla^2 f$ BE LIPSCHITZ ON S , FOR $L < \infty$

IF $\nabla^2 f(x^*)$ WITH x^* MINIMIZER IS DP AND $\|x_0 - x^*\| \leq \epsilon$ THEN QUADRATIC CONVERGENCE

A LITTLE NOTE

JACOBIAN

$$J_f(x) = \nabla^T f(x) = \begin{bmatrix} f_1 & \dots & f_m \end{bmatrix} \begin{matrix} x_1 & \dots & x_m \end{matrix}$$

LINEAR EQ $f(x) = Ax - b \quad \nabla f(x) = A^T$
 $J_f(x) = A$

QUADRATIC FUNCTION $\nabla f(x) = Qx - b$
 $\nabla^2 f(x) = Q$

WEIRSTRASS

K COMPACT* SET WITH $f: K \rightarrow \mathbb{R}$ CONTINUOUS

$\Rightarrow \exists \sup f, \inf f \quad * \mathbb{R}^m$: CLOSED AND BOUNDED

CHAPTER 2: LINEAR PROGRAMMING

LP FORMS

CANONICAL

$$\min C^T X + 0$$

$$Ax \geq 0 \geq 0$$

$$x \geq 0$$

$$\begin{aligned} UB=K & \quad X'_1 = K - X \\ LB=K & \quad X = X - K \\ \text{FREE} & \quad X' = X_+ - X_- \end{aligned} \quad X' \geq 0$$

STANDARD

$$\min C^T X + 0$$

$$Ax = 0 \geq 0$$

$$x \geq 0$$

$$\begin{aligned} \text{SLACKS} & \leq D + S \\ & \geq S - S \\ \text{RHS MUST BE POS!} \end{aligned}$$

EXTREME POINT

$X \in S$ IS EXTREME IFF $\nexists Y, Z \in S$:
 $X = \lambda Y + (1-\lambda)Z, \lambda \in (0,1)$

RANK $A_m(x) = m$ m LIN. INDEP. ROWS

BASIC SOLUTION

"GIVEN $A^{m \times m}$, WITH FULL RANK, THE BASE B IS A SUBMATRIX OF A WITH m LINEARLY INDEPENDENT COLUMNS"

$$A = [B | N] \quad \det(B) \neq 0 \quad x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}, x_B \in \mathbb{R}^m$$

FEASIBLE

$$x \geq 0$$

$$x = 0 \quad \text{DEGENERATE RED. CONSTR.}$$

BOUNDED BY A FACTOR (m)
 EXTREME IS ALSO BF.

ADJACENT

2 BASES SHARE $m-1$ CONSTRAINTS

MINKOWSKI'S WEYL

"A POLYHEDRON IS EXPRESSED AS A CONVEX COMBINATION OF EXTREME POINTS, AND EXTREME RAYS"

$$P := \left\{ \sum_i \lambda_i V_i + \sum_j \alpha_j V_j : \sum_i \lambda_i = 1, \lambda_i, \alpha_j \in \mathbb{R}^+ \right\}$$

THM

IF THERE IS AN OPTIMAL SOLUTION, THEN THERE IS AN OPTIMAL BASIC FEASIBLE SOLUTION

PROOF

$$MW \quad x = d + \sum_k \lambda_k V_k \quad Ad = 0 \quad d \geq 0$$

UNBOUND $C^T d \geq 0$ OTHERWISE

$$\begin{cases} x_\gamma = \gamma d + \sum_k \lambda_k V_k \\ C^T(x_\gamma) = \gamma C^T d \text{ UNBOUND} \end{cases}$$

NO RAY \Rightarrow OPTIMAL

SIMPLEX TABLEAU

BASE $\begin{pmatrix} (1) \\ (2) \\ \vdots \\ \end{pmatrix}$
BASE m $\begin{pmatrix} z \end{pmatrix}$

$x_1 \quad x_2 \quad \dots \quad s_1 \quad \dots \quad s_m \quad b$

SIMPLEX DICT.

$$x_{B_1} = b_{B_1} + a_{B_1} x_1 \dots$$

$$x_{B_2} = b_{B_2} + a_{B_2} x_1 \dots$$

\vdots

$$z = \sum_D x_{D\bar{n}} \text{ NON BASICS}$$

RATIO $\frac{+b_{Be}}{-a_{Be} \neq 0!}$

PICK SMALLEST OR
MINIMUM INDEX IF TIE!

EX

MIN

$$-x_1 \quad -2x_2 \quad -3x_3$$

S.T.

$$x_1$$

$$+s_1 = 2$$

$$x_2$$

$$+s_2 = 2$$

$$x_3$$

$$+s_3 = 2$$

$$x_1 + x_2 + x_3 + s_4 = 5$$

$$x_i \geq 0 \quad \forall i \in \{1, 2, 3, 4, 5\}$$

BASIC SOLUTION WITH SLACKS

$$s_1 = 2 - x_1$$

RATIO $2/-1 = 2$ WIN!

$$s_2 = 2 - x_2$$

$$2/0 = +\infty$$

$$s_3 = 2 - x_3$$

$$2/0 = +\infty$$

$$s_4 = 5 - x_1 - x_2 - x_3$$

$$5/1 = 5$$

$$z = -x_1 - 2x_2 - 3x_3$$

IN = x_1 OUT = s_1

SUBSTITUTE $x_1 = 2 - s_1$

$$x_1 = 2 - s_1$$

RATIO $2/0 = +\infty$

$$s_2 = 2 - x_2$$

$$2/1 = 2$$
 WIN!

$$s_3 = 2 - x_3$$

$$2/0 = +\infty$$

$$s_4 = 3 + s_1 - x_2 - x_3$$

$$3/1 = 3$$

$$z = -2 + s_1 - 2x_2 - 3x_3$$

IN = x_2 OUT = s_2

SUBSTITUTE $x_2 = 2 - s_2$

$$X_1 = 2 - S_1$$

RATIO

$+\infty$

$+\infty$

$$X_2 = 2 - S_2$$

$$S_3 = 2 - X_3$$

$$2/1 = 2$$

$$S_4 = 1 + S_1 + S_2 - X_3 \quad 1/1 = 1 \quad \text{WIN}$$

$$Z = -6 + S_1 + 2S_2 - 3X_3 \quad \text{OUT} = S_4 \quad \text{IN} = X_3$$

SUBSTITUTE

$$X_3 = 1 + S_1 + S_2 + S_4$$

$$X_1 = 2 - S_1$$

RATIO

$$2/1 = 2$$

$$X_2 = 2 - S_2$$

$+\infty$

$$S_3 = 1 - S_1 - S_2 + S_4 \quad 1/1 = 1 \quad \text{WIN}$$

$$X_3 = 1 + S_1 + S_2 - S_4 \quad 1/-1 = -1$$

$$Z = -9 - 2S_1 - S_2 + 3S_4 \quad \text{IN} = S_1 \quad \text{OUT} = S_3$$

SUBSTITUTE

$$S_1 = 1 - S_2 - S_3 + S_4$$

$$X_1 = 2 - S_1 = 1 + S_2 + S_3 - S_4$$

$$X_2 = 2 - S_2$$

$$S_1 = 1 - S_2 - S_3 + S_4$$

$$X_3 = 2 - S_3$$

$$Z = -11 + S_2 + 2S_3 + 2S_4 \quad \text{OPTIMAL}$$

SIMPLEX
MATRIX

A. ARTIFICIAL PROBLEM

SLACKS AS FIRST BASE

$Z > 0$: INFEASIBLE

B. COMPUTE STUFF LOOPING

1. INVERT $B \rightarrow B^{-1}$

2. MULTIPLICATORS

3. REDUCED COSTS

OUT-VARIABLE

$$K^T = C_B^T \cdot B^{-1} \quad X_B = B^{-1} \cdot b$$

$$K^D = C_D^T - K^T D \quad \text{IN: BLAND}$$

$$X_q = B^{-1} \cdot A_{\text{IN}} \quad \text{SMALLEST}$$

$$X_q < 0 \quad \text{UNBOUNDED}$$

$$C. Z = C_B^T X_B$$

INITIAL BASE

SLACK IF EVERY CONSTRAINT
HAS ONE!

THM IF ALL THE REDUCED COSTS OF A BASIC FEAS. SOL ARE POSITIVE, THEN THE SOLUTION IS OPTIMAL

PROOF ASSUME X B.F.S WITH $r \geq 0$

$$X_H = 0; X_B \geq 0 \Rightarrow z = C_B^T X_B$$

$$Y \text{ FEASIBLE} = (\hat{X}_B, \hat{X}_H)$$

$$z(Y) = \hat{C}_B^T \hat{B}^{-1} \cdot b + \hat{C}_H^T \hat{X}_H$$

$$= \hat{C}_B^T \hat{X}_B + \hat{C}_H^T \hat{X}_H$$

$$= z + \hat{C}_H^T \hat{X}_H$$

□

PHASE 1 SIMPLEX

A. FROM STANDARD
SET SLACKS AS BASIC AND SOLVE

B. FOR EACH SLACK

IF POSITIVE BASIC

IF NEGATIVE ADD A NEW SLACK AND MAKE IT BASIC

C. OBJECTIVE

MIN $\sum_k X_B$ BASIC VARIABLES + BIGH!?!?

ALTERNATIVELY

$$z_{P1} = \min_t t$$

$$Ax + te^T \geq b \quad e = [1 \dots 1]^T \in \mathbb{R}^m$$

$$x, t \geq 0$$

A. SLACK IN BASE

B. BRING t IN BASE PIVOT ON "MOST INFEAS"
SLACKS FEASIBLE?

$z > 0$ STOP. INFEASIBLE

$z = 0$ FEASIBLE BASE

THM z_{P1} IS FEASIBLE AND BOUNDED
 $0 \leq z_{P1} < \infty$

PROOF $(t, x) = (b_t, 0)$ IS FEASIBLE

$$\begin{cases} Ax + te = b_t e \geq b \\ z_{P1} \leq b_t < \infty \wedge z_{P1} \geq 0 \end{cases}$$

□

THM THE ORIGINAL PL IS FEASIBLE IFF $z_{P_1} = 0$

PROOF \Rightarrow IF PL IS FEASIBLE

$$\exists \hat{x} \geq 0 : A\hat{x} \geq b$$

BUT $(0, \hat{x})$ IS FEASIBLE FOR z_{P_1} WITH $z_{P_1} = 0$

\Leftarrow IF $z_{P_1} = 0$, $\exists (\hat{t}, \hat{x}) \geq 0 : A\hat{x} + \hat{t}e \geq b$
AND $z_{P_1} = \hat{t} = 0$

EX (PL) MIN $x_1 + x_2$

$$-x_1 + x_2 \geq 1$$

$$2x_1 + x_2 \geq 4$$

$$x \geq 0$$

$$\Rightarrow \begin{aligned} S_1 &= -1 + x_2 - x_1 \\ S_2 &= -4 + 2x_1 + x_2 \\ z &= x_1 + x_2 \end{aligned} \quad x = \begin{pmatrix} 0 \\ 0 \\ -1 \\ -4 \end{pmatrix} \text{ INFEAS!}$$

(z_{P_1}) MIN t MOST INFEAS

$$S_1 = -1 + x_2 - x_1 + t \quad \text{PIVOT ON } t$$

$$S_2 = -4 + 2x_1 + x_2 + t$$

$$z = t$$

$$S_1 = 3 - 3x_1 + S_2$$

$$t = 4 - 2x_1 - x_2 + S_2 \quad \text{PIVOT ON } S_2$$

$$z = 4 - 2x_1 - x_2 + S_2$$

$$S_1 = 3 - 3x_1 + S_2$$

$$x_2 = 4 - 2x_1 - t + S_2$$

$$z_A = t = 0 \quad \text{OPTIMAL!}$$

(PL) WITH BASE $x = (0 \ 4 \ 3 \ 0)$

$$S_1 = 3 - 3x_1 + S_2 \quad \text{WITHOUT } t!$$

$$x_2 = 4 - 2x_1 - \cancel{x} + S_2 \quad \begin{matrix} 3/3=1 \\ 4/2=2 \end{matrix}$$

$$z = 4 - x_1 + S_2 \quad \text{IN } x_1 \text{ OUTS!}$$

$$x_1 = 1 - 1/3 S_1 + 1/3 S_2$$

$$x_2 = 2 + 2/3 S_1 + 1/3 S_2$$

$$z = 3 + 1/3 S_1 + 2/3 S_2 \quad \text{OPTIMAL}$$

THE BLAND'S
RULE

GET IN THE BASE THE CANDIDATE WITH
THE SMALLEST INDEX, SAME FOR RATIO TEST

THM WITH GOLD OLD BLAND, THE SIMPLEX
TERMINATES

PROOF ASSUME BLAND + CYCLE $D_0 \dots D_r = D_0$
 $I = \{\text{SET OF VARS MOVING IN/OUT BASE}\}$
 $\rho = \max I$

$$\text{AT } D_r: \begin{aligned} x_r &= b_r - \sum_{s \notin D_r} a_{rs} x_s \\ z &= v + \sum_r c_r^* x_r \end{aligned} \quad \text{0 FOR } j \in D_r = B_r \text{ etc...}$$

WHO DOES WHAT

$$x_{IN} \Rightarrow \hat{c}_{IN} < 0$$

$$x_{OUT} \Rightarrow \hat{c}_{OUT} > 0$$

$$\text{BLAND} \quad \hat{a}_j \leq 0 \quad \forall j \neq \text{OUT}, j \in I \text{ IN BASE}$$

CONSIDER 2 SOLUTIONS

$$\hat{x} = \begin{cases} x_B = 0 \\ x_B = B^{-1}b \\ z(\hat{x}) = \hat{z} \end{cases} \quad \hat{y} = \begin{cases} y_H = 0 \text{ BUT } y_i = 1 \\ y_B = x_B - a_i \bar{x} \\ z(\hat{y}) = \hat{z} + C_i \bar{x} < z(\hat{x}) \end{cases}$$

$$z(\hat{x}) > z(\hat{y}) \Rightarrow r_N^T x_N > r_N^T x_N$$

$$r^T x > r^T y$$

$$r_B^T \hat{a} > r_{\bar{x}} \geq 0 \quad \text{NO LOOP!} \quad \square$$

CHAPTER 3: ADVANCED LP

DUALITY

PRIMAL

$$\text{MIN } C^T X$$

$$= b^T \bar{z}$$

$$X_i \geq 0$$

$$X_i \text{ FREE}$$

$$\text{MAX} \rightarrow \text{MIN}$$

THM

THE DUAL OF THE DUAL IS THE PRIMAL

PROOF

$$(P) \text{ MIN } C^T X \text{ s.t. } Ax \geq b, x \geq 0$$

$$(D) \text{ MAX } b^T y \text{ s.t. } A^T y \leq C, y \geq 0$$

GO TO CANONICAL

$$(D') \text{ MIN } -b^T y \text{ s.t. } -A^T y \geq -C, y \geq 0$$

$$(D')^D \text{ MAX } -C^T x \text{ s.t. } -A^T x \leq -b, x \geq 0$$

NOTE BETTER TO REDUCE CONSTRAINTS!

STORAGE

$$(m+1)(m+1) \quad (m+1)(m+2)$$

$$\text{DENSE } -m^2 + mn + m + 1 \quad mn + (m+1)^2$$

PIVOTS

$$\text{SPARSE } O(m(m-m)) \quad O(m^2)$$

DICT

REVISED

WEAK DUALITY

THM x, y FEASIBLE FOR PRIMAL AND DUAL, THEN $C^T x \geq b^T y$

PROOF DUAL CONSTRAINTS $C^T \geq A^T y \quad x \geq 0 \Rightarrow$



$$C^T x \geq (A^T y) x = y^T A x = y^T b$$

$$C^T x \geq b^T y$$

COROLLARY 1 (P) UNBOUNDED \Rightarrow (D) INFEASIBLE
(D) UNBOUNDED \Rightarrow (P) INFEASIBLE
OR BOTH INFEASIBLE / FEASIBLE *

COROLLARY 2

$$\text{OPTIMUM IS } C^T x = b^T y$$

BOTH INFEAS: $C^T x \neq b^T y$!!

STRONG DUALITY

THM IF ONE LP, EITHER PRIMAL OR DUAL, HAS AN OPTIMUM, SO DOES THE OTHER ONE

PROOF ASSUME $X = (X_B \ X_D)^T$ OPTIMAL

RED. COSTS $C_D^T - h^T D \geq 0$
 $y^T = h^T$

FEASIBLE $y^T A = C_B^T B^{-1} (B \ D)$
 $= C_B + y^T D \leq C^T$

OPTIMAL (P) $z = C^T X = C_B^T (B^{-1} b)$
 (D) $w = b^T y = C_B^T B^{-1} b = z$

□

DUAL SOL $h^T = y^T = C_B^T \cdot B^{-1}$ PRIMAL SHADOW

DUAL SLACK $v = c - A^T y$ REDUCED COSTS!
 PRIMAL

COMPLEM. SLACKNESS

THM x^*, y^* ARE PRIMAL-DUAL OPTIMAL IFF

$(A^T y - c)^T x^* = y^{*T} (b - A x^*)$

(D) CONST (P) SOL (D) (P) CONST

ALTERNATE VERSION $\forall i: x_i > 0 \Rightarrow y_i^T A_i = c_i$
 $\forall j: y_j > 0 \Rightarrow (A x)_j = b_j$

PROOF $(y^T A) x \leq C^T x^* = y^T b \leq y^T (A x)$
 $\Rightarrow (y^T A - C^T) x^* = 0 = y^T (A x^* - b)$ □

VERIFY OPTIMALITY

$\exists y \in \mathbb{R}^m:$

$$\begin{cases} (y^T A)_j = c_j & \text{IF } x_j > 0 \\ y_i = 0 & \text{IF } (A x)_i \geq b_i \\ y \geq 0 & y^T A \leq C^T \end{cases}$$
 FEASIBILITY

DEGEN. (INF. SOL) IF P HAS AN OPTIMAL DEGENERATE, D HAS A SINGLE OPTIMUM. AND VICES VERSA. DOES NOT IMPLY UNIQUE \Leftrightarrow MULTIPLE FOR ALL PROBLEMS!

FARKAS LEMMA

// LET $A \in \mathbb{R}^{m \times n}$ $b \in \mathbb{R}^m$. THEN EITHER ONE OF THE FOLLOWING IS TRUE

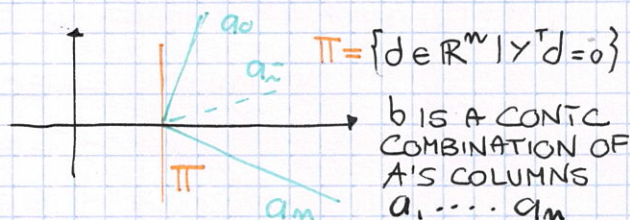
A. $\exists x \in \mathbb{R}^n : Ax = b \wedge x \geq 0$

B. $\exists y \in \mathbb{R}^m : (A^T y \geq 0 \wedge b^T y < 0)$ //

INTERPRET. CONE $C = \{ \sum_{i=1}^n \lambda_i a_i \mid \lambda_i \geq 0 \}$ COLUMNS OF A

IF A $b \in C$ AND ALL a_i ARE ON "THE SAME SIDE" OF Π

IF B $b \notin C$ LINEAR SEPARATOR



PROOF DUALITY USE IT!!!

(P) = MAX 0 S.T. $Ax = b, x \geq 0$

(D) = MIN $b^T y$ S.T. $A^T y \geq 0$ \square

$y^T Ax = 0 \quad y^T b \neq 0$

DEPENDENT ON MIN/MAX

SENSITIVITY ANALYSIS

VARIATION OF b

FEASIBILITY.

b

NEED TO CHECK THE FEASIBILITY

$x_B = B^{-1}b \geq 0$ WITH $b' = b + \theta$

$\begin{cases} \Delta z = \overline{C^T \cdot \theta} = C^T \Delta x \\ x_B = B^{-1}b' \geq 0 \end{cases}$ DUAL IS BETTER

THM $V: b \in \mathbb{R}^m \rightarrow \min_x C^T x : Ax = b, x \geq 0$

IS A CONVEX FUNCTION

PROOF LET $0 \leq \alpha \leq 1, b_1, b_2 \in \mathbb{R}^m$
IF $V(b_1), V(b_2) = \infty$ \square
OTHERWISE

x_1, x_2 SOLUTIONS OF P_1, P_2

$x_0 = \alpha x_1 + (1-\alpha)x_2 \Rightarrow$

$Ax_0 = \alpha Ax_1 + (1-\alpha)Ax_2$

$V(\alpha b_1 + (1-\alpha)b_2) = \min_x C^T x$
 $Ax = \alpha b_1 + (1-\alpha)b_2$

$\leq C^T x_0$ \square

SHADOW PRICE NEGATIVE PRIMAL REDUCED COST FOR THE VARIABLE (SACR)

FUNDAMENTAL THM OF SENSITIVITY

THM $b \in \mathbb{R}^m$: $v(b) = \{ \min C^T x : Ax = b, x \geq 0 \}$
 IS A **FINITE VALUE**. IF THE
DUAL SET $\mathcal{L}(b)$ IS **NON EMPTY**

$$\mathcal{L}(b) = \{ y \in \mathbb{R}^m : y^T A \leq C, y^T b = v(b) \}$$

AND **BOUNDED**, THEN

$$\forall d \in \mathbb{R}^m, v'_d(b) = \max_{y \in \mathcal{L}(b)} y^T d$$

COROL IF $|\mathcal{L}(b)| = 1$, THEN THE
MARGINAL CONTRIBUTION OF 1TH
CONSTRAINT IS ITS RED. COST

OPTIMALITY CHECK THE SOLUTION IS OPTIMAL

$$\begin{cases} C = C + \theta & K^T = C_B^T B^{-1} \\ K_D = C_D^T - K^T D \geq 0 \end{cases}$$

**BOUNDED
SIMPLEX**

REPLACING IS INEFFICIENT

**BASE
ENTER**

$$X_i \neq LB_i \neq UB_i$$

$$\begin{cases} X_i = LB_i \wedge K_i \leq 0 \\ X_i = UB_i \wedge K_i > 0 \end{cases}$$

**RATIO
TEST?**

A. $m \in$

$$\text{B. } \begin{cases} \min \frac{b_i}{a_{i,t}} : a_{i,t} > 0 & a_{i,t} < 0 \\ \min \frac{m_i - b_i}{a_{i,t}} : a_{i,t} < 0 & a_{i,t} > 0 \end{cases}$$

$$\begin{aligned} // X_i &= UB_i \\ // X_i &= 0 \end{aligned}$$

**DUAL
SIMPLEX**

EXIT

MOST INFEAS

ENTERS

$$\min \left| \frac{C_j}{a_{s,j}} \right|, a_{s,j} < 0 \quad (\text{HAS - IN TABLEAU})$$

$$\begin{aligned} X_1 &= 8/5 & -2/5 X_3 + 1/5 X_6 \\ X_4 &= -11/5 & + X_2 + 9/5 X_3 - 2/5 X_6 \\ X_5 &= -9/5 & + 4 X_2 + 1/5 X_3 - 3/5 X_6 \\ Z &= -32/5 & - 8 X_2 - 37/5 X_3 - 1/5 X_6 \end{aligned}$$

$$\begin{aligned} X_4 &\text{ LEAVE} & X_2 &= 8/1 & X_3 &= \frac{37}{5} \cdot \frac{5}{9} \\ X_3 &\text{ ENTERS} & & & & \end{aligned}$$

CHAPTER 4: UNCONSTRAINED NON-LINEAR (AND NOT!)

EPHIGRAPH

$f: \mathcal{X} \rightarrow \mathbb{R}$ IS CONVEX IF ITS EPHIGRAPH
 $\{(x, z) \in \mathbb{R}^m \times \mathbb{R} : z \geq f(x) \forall x \in \mathcal{X}\}$ IS
 CONVEX

EX $f(x) = x^2$

$\forall x, y \in \mathbb{R}, \alpha \in [0, 1]$

$$\begin{aligned} f(\alpha x + (1-\alpha)y) &= \alpha^2 x^2 + 2\alpha(1-\alpha)xy + (1-\alpha)^2 y^2 \\ &\leq \alpha^2 x^2 + 2\alpha(1-\alpha)(x^2 + y^2) + (1-\alpha)^2 y^2 \\ &= \alpha^2 x^2 + 2\alpha(1-\alpha)(x^2 + y^2) + (1-\alpha)^2 y^2 \end{aligned}$$

THE LATER IS CONVEX

TAYLOR²

UNIVAR $f: \mathbb{R}^m \rightarrow \mathbb{R}^m \quad F(x+d) = F(x) + J_F(x) \cdot d$

MULTIVAR $g: \mathbb{R}^m \rightarrow \mathbb{R} \quad G(x+d) = G(x) + J_G(x) \cdot d + \frac{1}{2} d^T H_G(x) d$
 $z \in [0, 1]$

UNDER ESTIMATOR IF f CONVEX AND C^1 , THEN $T_f^{(1)}$ IS AN UNDERESTIMATOR

OPTIMALITY CONDITIONS

$f \in C^2$:

1ST

2ND

MIN LOCAL
MAX

$$\nabla f = 0$$

$\nabla^2 f$ DP (SDP IS NEC)
DN

MIN GLOBAL
MIN

//

// $\forall x$
// $\forall x$

SADDLE

$$\nabla f = 0$$

$\nabla^2 f$ NOT (DPV DN) !S!
 $\exists d^+ d^-:$
 $d^+ T \nabla^2 f d^+ < 0$
 $d^- T \nabla^2 f d^- > 0$

???

$$\nabla f = 0$$

$\nabla^2 f$ NOT (DPV DN)

NECESSARY + SUFFICIENT DP(DN)
NECESSARY SDP(DN)

IF SDP + (?) NOT SUFFICIENT!

C^k + POINT \Rightarrow NECESSARY 1ST $C^1 \Rightarrow \nabla f = 0$

2ND $C^2 \Rightarrow \nabla^2 f$ SDP
SDN

C^k + $\nabla + \nabla^2$ \Rightarrow STRICT SUFFICIENT $C^2 \Rightarrow \nabla f = 0$

$C^2 \Rightarrow \nabla^2 f$ DP

A FEW EXAMPLES

QUADRATIC OPTIMIZATION

$$\min_{x \in \mathbb{R}^m} f(x) = \frac{1}{2} x^T Q x + c^T x + d \quad Q \text{ SYMMETRIC}$$

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

$$\nabla^2 f(x) = Q \quad \text{d/ROW} = \text{COL}$$

$$1^{\text{ST}} \quad Qx = -c \quad 2^{\text{ND}} \quad Q(S)DP$$

L2 NORM

$$\min_{x \in \mathbb{R}^m} \|Ax - b\|_2^2 \quad A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, \rho(A) = m \leq n$$

$$= x^T (A^T A) x - 2x^T A^T b + b^T b$$

$$1^{\text{ST}} \quad \nabla f(x) = 2(A^T A)x - 2A^T b = 0$$

$$x = (A^T A)^{-1} A^T b$$

PSEUDO-INVERSE

$$2^{\text{ND}} \quad \nabla^2 f(x) = 2A^T A = 0 \quad DP!$$

EQUALITY CONSTRAINTS

$$\text{IDEA} \quad \min_{x \in \mathbb{R}^m} f(x) : h(x) = 0$$

CONTOUR LINES SO THAT $f(x) = d$ (CANNOT DECREASE) AND $h(x) = 0$ SAME RESCALED GRADIENTS

$$\nabla f = \lambda \nabla h$$

IN GENERAL

$$\nabla f(x) = \sum_{\bar{i}} \lambda_{\bar{i}} \nabla h_{\bar{i}}(x)$$

LOCAL MINIMUM

$$\exists \varepsilon > 0 : \forall x \in B(x^*, \varepsilon) :$$

$$h(x) = 0, f(x) \geq f(x^*) \quad \forall x$$

FIRST ORDER NECESSARY IF LICQ

LET x^* BE A LOCAL MINIMIZER OF f WITH $f \in C^1$ AND $h \in C^1 : h(x) = 0$. x_1, \dots, x_m

$$\left\{ \nabla h_{\bar{i}}(x^*) \mid \bar{i} \in [1, \dots, m] \right\} \begin{matrix} h_1 \\ \vdots \\ h_m \end{matrix}$$

FORMS AN INDEPENDENT SET, THEN

$$\exists ! \lambda \in \mathbb{R}^m : \nabla f(x^*) = \nabla h(x^*) \cdot \lambda$$

$$= (\nabla h_1 \dots \nabla h_m) \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}$$

$$\sum \lambda_{\bar{i}} \nabla h_{\bar{i}}(x^*) = 0 \Leftrightarrow \lambda_{\bar{i}} = 0 \quad \forall \bar{i}$$

EX $\min_{x \in \mathbb{R}^2} x - e^{x_1+x_2} \text{ s.t. } x_1^2 + x_2^2 = 1$

$$\nabla f = \begin{pmatrix} 1 - e^{x_1+x_2} \\ -e^{x_1+x_2} \end{pmatrix}$$

$$\nabla h = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} \Rightarrow \begin{cases} 1 - e^{x_1+x_2} = \lambda 2x_1 \\ -e^{x_1+x_2} = \lambda 2x_2 \\ x_1^2 + x_2^2 = 1 \end{cases}$$

EQUAL LAGRANGIAN

LAGRANGIAN PROBLEM

$$\min_{x \in \mathbb{R}^m, \lambda \in \mathbb{R}^m} L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i h_i(x) + \lambda_0 (c - h_0(x))$$

1ST ORDER NECESSARY

$$\nabla L = 0 = \begin{cases} \nabla_x L = \nabla f(x) - \lambda \nabla h(x) = 0 \\ \nabla_\lambda L = h(x) = 0 \end{cases}$$

TANGENT CONE

THE CONE TANGENT TO THE SET $\mathcal{A} = \{x \in \mathbb{R}^m : h(x) = 0\}$ IN $x^* \in \mathcal{A}$ IS

$$T(x^*) = \{d \in \mathbb{R}^m : \nabla h_j^T(x^*) d = 0, \forall j \in \mathcal{M}\}$$

EX $\nabla h(x^*) = \begin{pmatrix} 1.2 & 0 & 4 \\ 1.3 & 9 & 0 \end{pmatrix}$ FEASIBLE DIR AT x^* CONE IFF LIQP

$$d \in T(x^*) \text{ IF } d = Bv, v \in \mathbb{R}$$

2ND ORDER

NECESSARY

x^* SATISFIES THE 1ST NECESSARY. THEN, $\nabla_{xx}^2 L(x^*, \lambda)$ HAS TO BE SDP ON THE TANGENT CONE

$$d^T \nabla_{xx}^2 L(x^*, \tilde{\lambda}) d \geq 0 \quad \forall d \in T(x^*) \text{ SDP}$$

SUFFICIENT

$x^* \in \mathbb{R}^m, \lambda^* \in \mathbb{R}^m : h(x^*) = 0, \nabla L(x^*, \lambda^*) = 0$ AND

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*) d > 0 \quad \forall d \neq 0 \in T(x^*) \text{ DP}$$

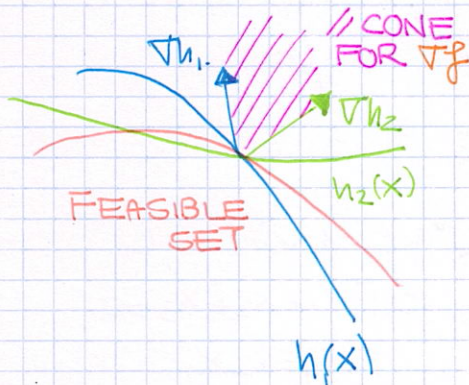
x^* IS STRICT MINIMUM!

\Leftrightarrow

$$z^T \nabla^2 f(x^*) z$$

WHERE $z = \text{NULL}(\nabla h_{\text{ACT}}(x^*))$

IF $\nabla g, h$ LIQP UNDER REGULARITY CONDITIONS



EX

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & \quad X_1 X_2 + X_1 X_3 + X_2 X_3 \\ \text{s.t.} & \quad X_1 + X_2 + X_3 = 3 \\ & \quad X_1^2 + X_n = 2 \end{aligned}$$

1ST ORDER

$$\begin{aligned} L(x, h) = & X_1 X_2 + X_1 X_3 + X_2 X_3 + \\ & -h_1 (X_1 + X_2 + X_3 - 3) - h_2 (X_1^2 + X_n - 2) \end{aligned}$$

$$\nabla L(x, h) = \begin{pmatrix} X_1 X_2 + X_3 - h_1 - 2h_2 X_1 \\ X_2 X_1 + X_3 - h_1 \\ X_3 X_1 + X_2 - h_1 \\ X_n - h_2 \\ h_1 X_1 + X_2 + X_3 - 3 \\ h_2 X_1^2 + X_n - 2 \end{pmatrix} \quad x^* = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\nabla L(x, h) = 0 \quad \text{IF } x = x^*, h^* = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \\ \text{WITH } f(x^*) = 3$$

2ND ORDER

$$\nabla^2 L_{xx}(x, h) = \begin{pmatrix} X_1 & X_2 & X_3 & X_n \\ X_1 & -2h_2 & 1 & 1 & 0 \\ X_2 & 1 & 0 & 1 & 0 \\ X_3 & 1 & 1 & 0 & 0 \\ X_n & 0 & 0 & 0 & 0 \end{pmatrix}$$

// = 0 WITH $h_2 = 0$

$$\nabla h^T(x^*) = \begin{pmatrix} h_1 & 1 & 1 & 0 & 0 \\ h_2 & 2 & 0 & 0 & 1 \end{pmatrix}$$

$$\begin{aligned} T(\hat{x}) = & \{ d \in \mathbb{R}^m \mid (1 \ 1 \ 1 \ 0) d = 0, (2 \ 0 \ 0 \ 1) d = 0 \} \\ & \begin{cases} X_1 + X_2 + X_3 = 0 \\ 2X_1 + X_n = 0 \end{cases} \quad \begin{cases} X_1 + X_2 + X_3 = 0 \\ X_n = -2X_1 \end{cases} \end{aligned}$$

SOLUTION

$$B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ -2 & -2 \end{pmatrix} B^T \nabla_{xx}^2 L(x^*, \lambda^*) B$$

$$\begin{pmatrix} 1 & -1 & 0 & -2 \\ 1 & 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ -2 & -2 \end{pmatrix} =$$

$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ -2 & -2 \end{pmatrix} = \begin{pmatrix} -2 & -1 \\ -1 & -2 \end{pmatrix} \begin{matrix} -2 \\ +5 \end{matrix}$$

$$\begin{pmatrix} +2 & +1 \\ +1 & +2 \end{pmatrix} \begin{matrix} DP \Rightarrow DN \\ \bar{x} \text{ MAX LOCAL} \end{matrix}$$

WITH (KKT) INEQUALITIES

LAGRANGIAN

$$L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i h_i(x) - \sum_{j=1}^K \mu_j g_j(x)$$

$$g_j(x) \geq 0 \quad - \text{MINIMIZING}$$

$$g_j(x) \leq 0 \quad + \text{SAME SIGN AS VIOLATION}$$

1ST ORDER NECESSARY

LET x^* BE A LOCAL MINIMUM OF $f \in C^1$ OVER $g(x) \geq 0$. IF

$\{\nabla g_i(x^*) \mid g_i(x^*) = 0\}$ FORM A LIN-INDEP, THEN

$$\exists! m \in \mathbb{R}^K, \begin{cases} \nabla f(x^*) = \nabla g(x^*) m \\ g_i(x^*) m = 0 \\ g(x^*) \geq 0 \quad m \geq 0 \quad (m = \mu) \end{cases} \text{COMPLEMENTARITY}$$

KKT CONDITIONS

1ST ORDER NECESSARY

KKT

WHEN BOTH $g(x)$ AND $h(x)$ LET x^* BE A LOCAL MINIMUM OF $f \in C^1$ ON $h(x) = 0, g(x) \geq 0$ AND $h, g \in C^1$. IF

$\{\nabla h_i(x) : i \in [m]\} \cup \{\nabla g_j(x) = 0 : j \in K\}$ IS LIN-INDEP THEN

$\exists h \in \mathbb{R}^m, \mu \in \mathbb{R}^K$ UNIQUE:

$$\nabla_x L(x^*, h, \mu) = 0 \quad \text{GRAD. OF } L \text{ IS COMB OF } h \text{ AND } g$$

$$h(\hat{x}) = 0, g(\hat{x}) \geq 0 \quad \text{PRIMAL FEAS}$$

$$g^T(x) \mu = 0 \quad \text{COMPLEMENTARITY}$$

$$\mu \geq 0 \quad \text{DUAL FEAS (THE DESCENT IS NOT FEAS!)}$$

$$k=0 \Rightarrow \text{INACTIVE}$$

INTERP

$L(x, h, \mu)$ HAS AN OPTIMAL SADDLE POINT MIN OVER x , AND MAX OVER h, μ

CONSTRAINT QUALIFICAT

REQUIRED FOR ALL OF THIS TO HOLD. LIQP: $\Rightarrow T(x) \approx f(x)$

MULTIPLIERS

$$Z' = Z + \lambda_p \cdot \epsilon$$

WITH SOME SIGN REASONING. SHADOW PRICE

QUADRATIC FORM OVER A SPHERE

SYMMETRIC STUFF

SYMMETRIC MATRIX

A SYMMETRIC $m \times m$

$$A = O \Lambda O^T \quad O \text{ ORTHOGONAL EIGENVECTORS}$$

A HAS m REAL EIGENVALUES! Λ DIAGONAL EIGENVALUES

$x^T A x$ IS NEGATIVE?

$$A = O \Lambda O^T \quad y = O^T x$$

$$\begin{aligned} x^T A x &= x^T O \Lambda O^T x = y^T \Lambda y \\ &= \sum_{i=1}^m \lambda_i \cdot y_i^2 \end{aligned}$$

IFF $\lambda_i \geq 0 \forall i$ THEN IT IS ALWAYS POSITIVE

ELLIPSOIDS

B PSD MATRIX $x^T B x \leq b$

$$B = O \Lambda O^T \quad y = O^T x$$

$$x^T B x \leq b \Rightarrow x^T O \Lambda O^T x \leq b$$

$$y^T \Lambda y \leq b \Rightarrow \sum_{i=1}^m \lambda_i y_i^2 \leq b$$

$$\begin{array}{ll} \text{MAX} & x^T A x \\ \text{s.t.} & x^T x = 1 \end{array} \quad \begin{array}{l} A \text{ SYMMETRIC} \\ \text{OR } x^T B x = 1 \quad B \text{ PSD} \end{array}$$

$$L(x, \lambda) = \lambda^T A x - \lambda x$$

$Ax = \lambda x$ EIGENVALUES + VEC!

$x^T x = 1$ IFF $x_i = x_j$ SAME EIGV.

SYMMETRIC PD B

$$B^{1/2} = O \Lambda^{1/2} O^T \quad \Lambda^{1/2} = (\sqrt{\lambda_1} \dots \sqrt{\lambda_m})$$

$$B^{-1/2} B B^{-1/2} = I \quad x = B^{-1/2} y$$

$$x^T B x = y^T B^{-1/2} B B^{-1/2} y = y^T y$$

$$y^T B^{-1/2} A B^{-1/2} y$$

SAME

CHAPTER "I MESSED UP WITH THE NUMBERING" (5)

LINE-SEARCH METHODS

"GIVEN A DESCENT DIRECTION d_k , TRY TO DETERMINE $\alpha_k > 0$ SO THAT

$$x_{k+1} = x_k + \alpha_k d_k$$

$$f(x_{k+1}) < f(x_k) \quad "$$

EXACT

$$\min_{\alpha_k} f(\alpha_k) = f(x_k + \alpha_k d_k)$$

↑ HEY, I'M EXPENSIVE!

ARMIJO'S CONDITION

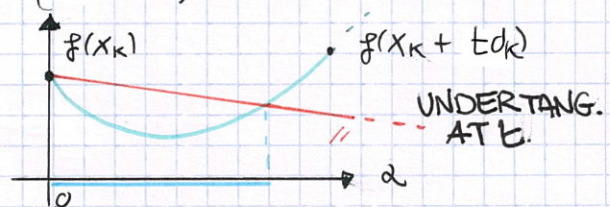
DECREASE WITHOUT TOO LONG STEPS

SUFFICIENT

INTUITION

FROM TAYLOR APPROXIMATIONS, ASSUMING WE HAVE A SIGNIFICANT DESCENT DIRECTION $d \in (0, 1)$

$$\begin{cases} f(x_k + t d_k) \leq f(x_k) + \underline{c}_x t d_k^T \nabla f(x_k) \\ \underline{c}_x \in (0, 1) \end{cases}$$



// VALID ARMIJO $\alpha = \mu \cdot \alpha_k$

THM LET $\alpha \in (0, 1)$ A FIXED PARAM. $\underline{c} > 0$ SATISFY THE ARMIJO IFF

$$f(x_k + t d_k) \leq f(x_k) + \alpha \underline{c} d_k^T \nabla f(x_k)$$

BACKTRACK START WITH LARGE t AND RESCALE IT AS

$$\underline{c}_{k+1} = \underline{c}_k \cdot \beta, \quad \beta \in (0, 1)$$

WOLFE'S CONDITION

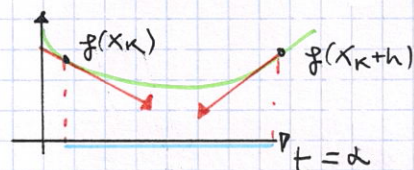
PREVENTS SHORT STEPS

CURVATURE

INTUITION

THE GRADIENT IN $f(x_k)$ SHOULD BE \underline{c}_2 TIMES GREATER THAN THE ONE IN $f(x_k + \alpha_k d_k)$

$$|d_k^T \nabla f(x_k + t d_k)| \geq \underline{c}_2 |d_k^T \nabla f(x_k)|$$



// VALID FOR WOLFE
"CURVATURE CONDITION"
SLOPE OF $f(x_k + \alpha t)$ IS \underline{c}_2 TIMES THE ONE OF $f(x_k)$

THM LET $C_2 \in (C_1, 1)$ WITH d FROM ARMIJO.
 $t > 0$ SATISFIES WOLFE IFF

$$d_k^T \nabla f(x_k + t d_k) \geq C_2 d_k^T \nabla f(x_k)$$

FREE LUNCH 1. ∇f L-LIPSCHITZ

$$2. 0 < C_1 < C_2 < 1$$

$$3. \theta = - \frac{\nabla f^T(x_k) d_k}{L \|d_k\|^2}$$

4. C_2 IS WOLFE

$\Rightarrow C_2 \in I$ IS ARMIJO

$$I = [\theta(1-\beta), 2\theta(1-\alpha)]$$

GLOBAL CONVERGENCE

THM IF $\|\nabla f(x_k)\| \rightarrow 0$ WITH $k \rightarrow +\infty$
 THEN THE CONVERGENCE IS SAID
 TO BE GLOBAL

* NOT TO A GLOBAL
 EXTREME POINT!

THM LET $f \in C^1$ BE A LOWER BOUNDED
 FN. LIPSCHITZ OF L WITH A SEQ
 $\{x_k\}$ OF ARMIJO-WOLFE POINTS. THEN

$$\nabla f^T(x_n) \cdot \frac{d_n}{\|d_n\|} = -\|\nabla f(x_n)\| \rightarrow 0$$

FREE LUNCH GOOD CONDITIONED ARMIJO-W
 CONVERGES! IF NEWTON,
 $\nabla^2 f(x)$ NOT SINGULAR!

$$\cos \theta_k = - \frac{\nabla f^T(x_k) \cdot d_k}{\|\nabla f(x_k)\|}$$

QUASI NEWTON

IDEA $f(x_k + t d_k) = f(x_k) + \nabla f^T(x_k) d + \frac{1}{2} d^T B_k d$
 B_k APPROXIMATES $\nabla^2 f(x_k)$ $O(m^2)$
 WITH 1ST ORD.

MODIFIED NEWTON ASSUME $\nabla^2 f(x_k)$ IS REPLACED BY

$$\nabla^2 f(x_k) = \nabla^2 f(x_k) + M_k \text{ DP}$$

$$h_1 \leq \dots \leq h_m \text{ OF } \nabla^2 f(x_k)$$

$$\text{LOOP } \begin{cases} \nabla^2 f = V^T \cdot \text{DIAG}(h) \cdot V \\ M_k = (\Sigma - h_1) \cdot V^T V = I \end{cases} \text{ IF } h_1 < \epsilon$$

$$\nabla^2 f = \nabla^2 f(x_k) + M_k$$

IDEA 2

SECANT METHOD

$$g'(x_{k+1}) = \frac{g(x_{k+1}) - g(x_k)}{x_{k+1} - x_k}$$

2 GIVEN POINTS!

$$\dots \text{BFGS } \nabla^2 f(x_{k+1}) \cdot (x_{k+1} - x_k) \approx \nabla f(x_{k+1}) - \nabla f(x_k)$$

NO VECTOR DIVISION \therefore

BFGS
NEWTON

EMULATE THE SECANT METHOD SO THAT: $\nabla^2 f_k \approx B_k$

1. $B_0 = I_n$

4. B_{k+1} DP

2. $B_{k+1}(x_{k+1} - x_k) = \nabla f_{k+1} - \nabla f_k$

5. B_{k+1} EASY-TO-INVERT

3. B_{k+1} SYMMETRIC

$$y_k = \nabla f(x_{k+1}) - \nabla f(x_k) \quad | \quad s_k = x_{k+1} - x_k$$

$$B_{k+1} = \frac{(y_k + B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k} + B_k$$

WITH ARMIJO
WOLFE

DEFINED POSITIVE AND EVERYONE'S
HAPPY

FOR $k=0, 1, \dots$

A. IF $x_k = \text{OPTIMAL}$ STOP

B. $B_k p_k = -\nabla f(x_k)$ FOR p_k
ARMJO FOR α

C. UPDATE B_{k+1}

TRUST
REGIONS

// DECIDE A REGION OF TRUTHFULNESS FOR
THE MODEL, AND PICK THE STEP LENGTH
IN THE REGION SUPERLINEAR WITH CG, DOGLE

DUAL TO LINE-SEARCH: (DIRECTION + STEP)
SINCE STEP THEN DIRECTION

$$m_k(p) = f(x_k) + \nabla f(x_k)^T p + \frac{1}{2} p^T B_k p$$

(MIN)

S.T. $\|p\| \leq \Delta_k$ RADIUS

IF $(B_k \text{ SDP} \wedge \|B_k^{-1} \nabla f(x_k)\| \leq \Delta_k)$
FULL STEP $p_k = -B_k^{-1} \nabla f(x_k)$

STEP
SIZE

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)}$$

ACTUAL REDUCT. / PREDICTED RED.

$\rho_k \approx 1$ EXPAND REGION! GOOD
RELIABILITY

$\rho_k \approx 0$ SHRINK REGION NO GOOD \therefore
 $\rho_k < 0$

SOLVING THE SUBPROBLEM

CAUCHY CLOSED APPROACH FOR COMPUTING P

$$P_K^c = -\gamma_K \frac{\Delta_K \cdot \nabla f(x_K)}{\|\nabla f(x_K)\|}$$
$$\gamma_K = \begin{cases} 1 \leq 0 \\ \min\left(\frac{\|\nabla f(x_K)\|}{\Delta_K \cdot \nabla f(x_K) B_K \nabla f(x_K)}, 1\right) \end{cases}$$

☹ THIS IS STEEPEST DESCENT
WITH A PARTICULAR STEP SIZE

DOGLEF FOLLOWS A V-SHAPED TRAJECTORY
INTERSECT REGION BOUNDARY AT MOST
ONCE!

CHOLESKY FACTORIZES B
EXACT

CG CONJUGATE GRADIENT SUPERLINEAR
AND FAST

! WARNING! NON-SMOOTH FUNCTION REDUCE
THE RADIUS BUT CAN USE
NONCONVEX MODELS