LECTURE 13: FEASIBLE DIRECTION METHODS

METHODS:

- UNCONSTRAINED METHODS
  - STEEPEST DESCENT
  - NEWTON'S METHOD
- METHODS FOR LPs
  - SIMPLEX
- METHODS FOR CONVEX PROBLEMS
  - SUBGRADIENT METHOD
- METHODS FOR ILP
- METHODS FOR CONSTRAINED PROBLEMS:
  1. FEASIBLE DIRECTION METHODS (TODAY)
  2. PENALTY METHODS (NEXT LECT.)

CONSIDER THE PROBLEM $f^* = \inf f(x)$

$$\text{s.t. } x \in X$$

WHERE X IS NONEMPTY, CLOSED & CONVEX, AND $f \in C^1$ ON X

- IDEA: GENERALIZE UNCONSTRAINED OPTIMIZATION METHODS

GENERAL FEASIBLE-DIRECTION DESCENT METHOD

0: INITIALIZE $x^0 \in X$. SET $k := 0$

1: FIND A FEASIBLE DESCENT DIRECTION $p_k \in \mathbb{R}^n$ SUCH THAT THERE EXISTS $\bar{\alpha} > 0$:

$$x_k + \alpha p_k \in X \quad , \forall \alpha \in [0, \bar{\alpha}]$$

$$f(x_k + \alpha p_k) < f(x_k), \quad \forall \alpha \in (0, \bar{\alpha}]$$

2: DETERMINE STEP LENGTH $\alpha_k > 0$: $x_k + \alpha_k p_k \in X$

$$f(x_k + \alpha_k p_k) < f(x_k)$$

3: UPDATE $x_{k+1} = x_k + \alpha_k p_k$

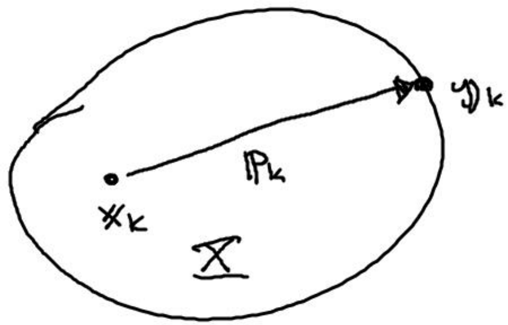
4: CHECK TERMINATION CRITERIA. OTHERWISE GOTO 1 AND LET $k = k+1$

- JUST AS <u>LOCAL</u> AS UNCONSTRAINED METHODS

- SEARCH DIRECTIONS ARE OFTEN ON THE FORM

  $p_k = y_k - x_k$ WHERE $y_k \in \underline{X}$ SOLVES SOME OTHER PROBLEM

- LINE SEARCH IS SIMILAR TO UNCONSTRAINED CASE. JUST DONT GO TO FAR!



- TERMINATION CRITERIA ARE OFTEN BASED ON KKT CONDITIONS

- FOR GENERAL X, FINDING FEASIBLE DESCENT DIRECTIONS CAN BE TRICKY. THE SAME HOLDS FOR STEP LENGTHS.

- IF X IS POLYHEDRON $\Rightarrow$ SEARCH DIRECTIONS AND STEP LENGTHS ARE EASY TO FIND

- IF X IS POLYHEDRON $\Rightarrow$ KKT CONDITIONS NECESSARY.
  x^* LOCAL MIN $\Rightarrow$ x^* KKT POINT

THE FRANK-WOLFE METHOD

- ASSUME $\min f(x)$ s.t. $x \in X$ WHERE X IS POLYHEDRON. AND $f \in C^1$ ON X

- OPTIMALITY CONDITIONS:

x^* LOCAL MIN OF f OVER $X \Rightarrow x^*$ STATIONARY

$$\Rightarrow \min_{x \in X} \nabla f(x^*)^T (x - x^*) = 0$$

FRANK-WOLFE

- IF $\min_{x \in X} \nabla f(x^*)^T (x - x^*) \neq 0 \Rightarrow x^*$ IS NOT LOCAL MIN

$\Rightarrow$ CAN CONSTRUCT FEASIBLE DESCENT DIRECTION AT x^*

- AT $x_k \in X$, IF

$$\begin{cases} \boxed{\min_{x \in X} \nabla f(x_k)^T (x - x_k)} < 0 \\ y_k \in \arg\min_{x \in X} \nabla f(x_k)^T (x - x_k) \end{cases}$$

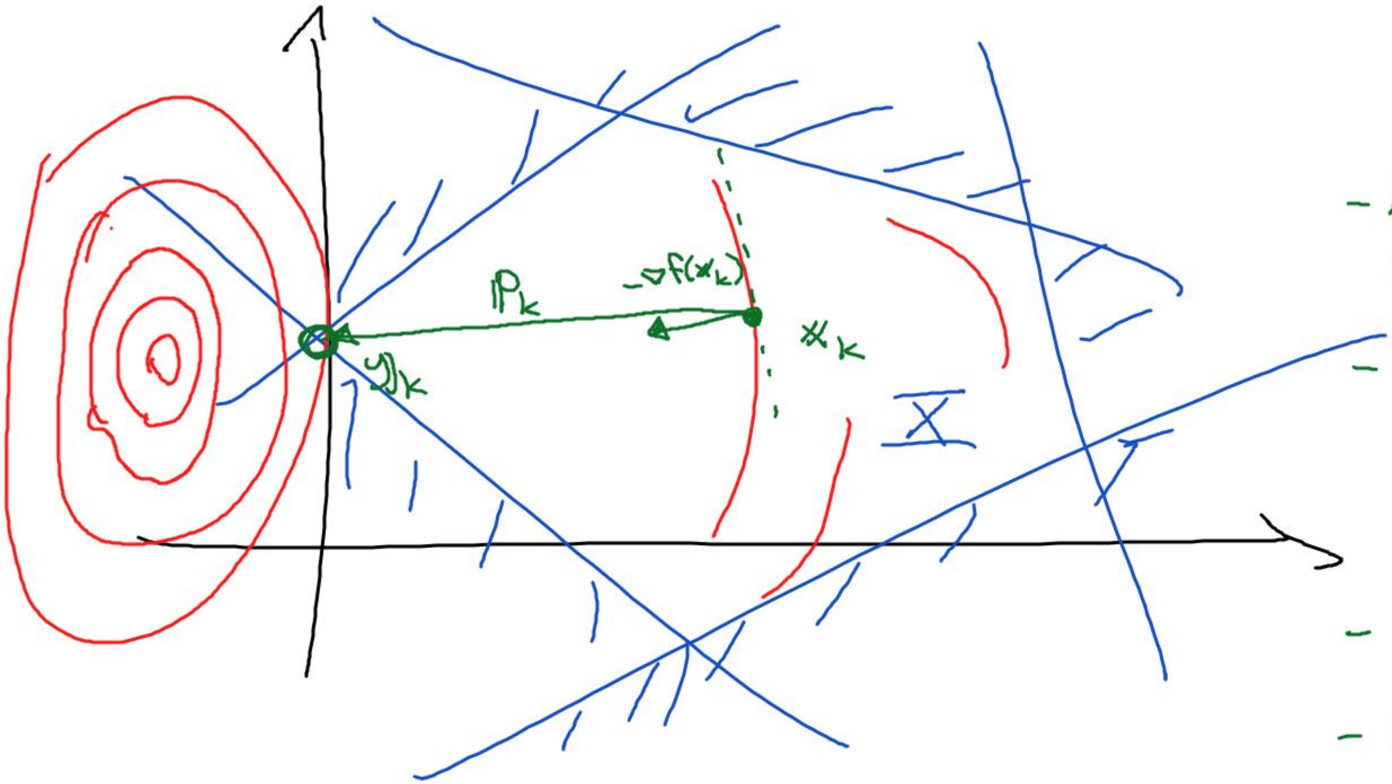
LINEAR PROGRAM!

- X POLYHEDRON
- OBJECTIVE LINEAR W.R.T. x
- EASY TO SOLVE

- THEN $p_k = y_k - x_k$ WILL BE A FEASIBLE DESCENT DIRECTION

- y_k WILL BE AN EXTREME POINT!

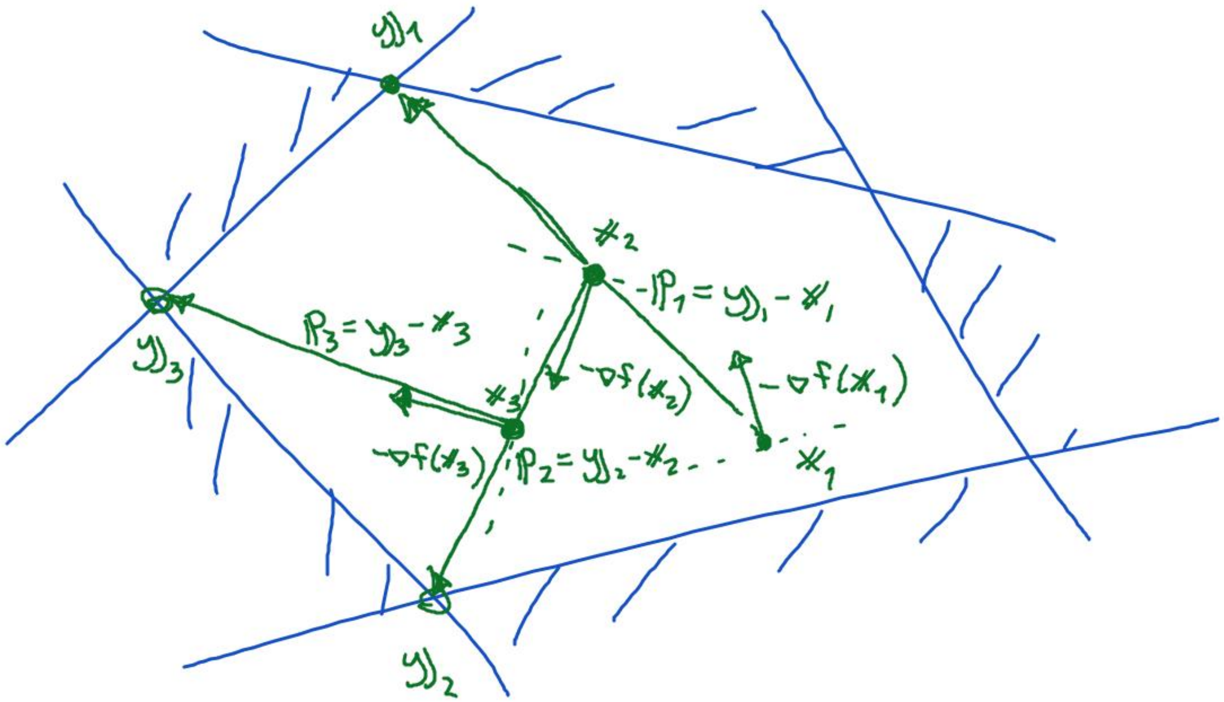
- SEARCH DIRECTION WILL ALWAYS BE TOWARDS AN EXTREME POINT IN THE FRANK-WOLFE METHOD



- AT ITERATION x_k WE COMPUTE $\nabla f(x_k)$.
- NOW SOLVE THE LP

$$\min \underbrace{\nabla f(x_k)^T}_{= c} x \quad \text{s.t.} \quad x \in X$$

- LET y_k BE SOLUTION
- LET SEARCH DIRECTION BE $p_k = y_k - x_k$



y_1
x_2
$-p_1 = y_1 - x_1$
$p_3 = y_3 - x_3$
y_3
x_3
$-\nabla f(x_2)$
$-\nabla f(x_1)$
$-\nabla f(x_3)$
$p_2 = y_2 - x_2$
x_1
y_2

FRANK-WOLFE METHOD

0: LET $x_0 \in X$. SET $k := 0$

1: FIND AN OPTIMAL SOLUTION TO THE PROBLEM (LP)

$$\min_{x \in X} \; z_k(x) = \nabla f(x_k)^T (x - x_k)$$

LET $p_k = y_k - x_k$

2: LINE SEARCH. FIND $\alpha_k \in [0, 1]$

WE WILL NEVER BECOME INFEASIBLE

3: $x_{k+1} = x_k + \alpha_k p_k$

4: IF $z_k(y_k)$ IS CLOSE TO 0 OR α_k IS CLOSE TO 0 WE TERMINATE, OTHERWISE LET $k = k+1$ AND GOTO 1.

CONVERGENCE:

- SUPPOSE $\underline{X}$ IS NONEMPTY POLYTOPE AND $f \in C^1$ ON $\underline{X}$.
  IF α_k IS CHOSEN ACCORDING TO ARMIJOS RULE THEN
  $\{x_k\}$ IS BOUNDED AND EVERY LIMIT POINT IS STATIONARY!

- IF f IS CONVEX ON $\underline{X}$, THEN EVERY LIMIT POINT IS
  GLOBAL OPTIMUM.

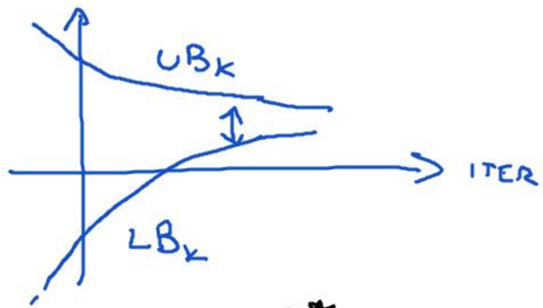
IF f IS A CONVEX FUNCTION

FOR EACH k, AND ALL $x \in X$ IT HOLDS THAT

$$f(x) \geq f(x_k) + \nabla f(x_k)^T (x - x_k)$$

[SINCE f IS CONVEX]

$$\geq \underbrace{f(x_k) + \nabla f(x_k)^T (y_k - x_k)}_{\text{GET THIS IN EACH ITERATION}}$$



$$\Rightarrow \quad f^* \geq f(x_k) + \nabla f(x_k)^T (y_k - x_k)$$

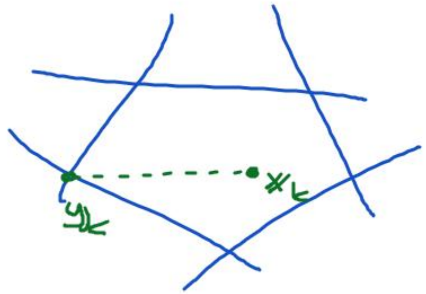
- SO IN EACH ITERATION WE GET A LOWER BOUND ON f^*
- KEEP BEST BOUND:

$$UB_{k+1} = \min \left[UB_k, f(x_k) \right] \quad \left[= f(x_k) \right]$$

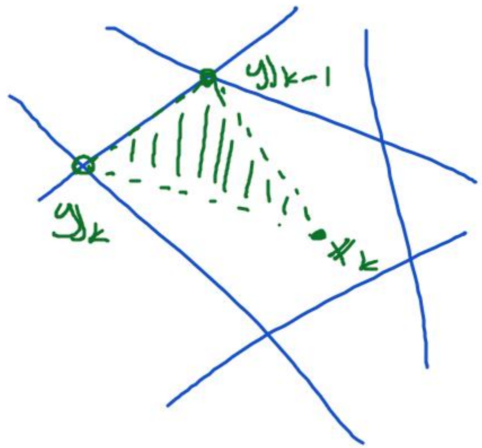
$$LB_{k+1} = \max \left[LB_k, f(x_k) + \nabla f(x_k)^T (y_k - x_k) \right]$$

SIMPLICIAL DECOMPOSITION

- FRANK-WOLFE METHOD FINDS NEXT ITERATE BETWEEN CURRENT ITERATE AND ONE EXTREME POINT
  - IN EACH ITERATION A NEW EXTREME POINT IS COMPUTED.



- SIMPLICIAL DECOMPOSITION METHOD REMEMBERS PREVIOUS EXTREME POINTS AND SEARCHES IN THE CONVEX HULL OF THESE AND THE CURRENT ITERATE FOR THE NEXT ITERATE

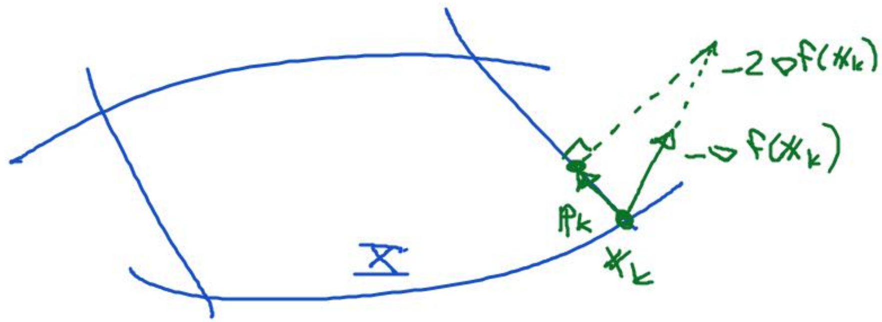


THE GRADIENT PROJECTION ALGORITHM

- IDEA: x^* LOCAL MIN $\Rightarrow$ $x^* = \text{Proj}_X \left[x^* - \nabla f(x^*) \right]$
- IF $x^* \neq \text{Proj}_X \left[x^* - \nabla f(x^*) \right]$ $\Rightarrow$ FEASIBLE DESC. DIR. CAN BE FOUND
- IF x_k IS NOT STATIONARY, THEN

$$p_k = \text{Proj}_X \left[x_k - \alpha \nabla f(x_k) \right] - x_k \quad \text{FOR SOME } \alpha > 0$$

IS A FEASIBLE DESCENT DIRECTION



- IF $X = \mathbb{R}^n$, THIS IS THE STEEPEST DESCENT METHOD

- HOW TO COMPUTE PROJECTIONS?
- ONLY POSSIBLE FOR SPECIFIC SETS
  - HYPERCUBE $\{x \in \mathbb{R}^n \mid 0 \leq x \leq 1\}$
  - UNIT SIMPLEX $\{x \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = 1, \ x \geq 0\}$
  - OTHER SPECIFIC TYPES OF SETS
- ANOTHER RELATED METHOD IS THE <u>GRADIENT PROJECTION METHOD WITH PROJECTION ARC</u>

$$x_{k+1} = \text{Proj}_{X}\left[x_k - \alpha_k \nabla f(x_k)\right]$$

- HOMEWORK: UNDERSTAND DIFFERENCE

CONVERGENCE

X IS NONEMPTY, CLOSED, CONVEX, AND $f \in C^1$ ON X.

IF α_k IS CHOSEN ACCORDING TO ARMIJOS RULE THEN

$\{x_k\}$ IS BOUNDED AND EVERY LIMIT POINT IS STATIONARY.

FRANK - WOLFE	GRADIENT PROJECTION METHOD
X POLYHEDRON	POSSIBLE TO PROJECT POINTS ON X