## **Duality**

#### How do we get an upper bound to a maximization LP?

max 
$$13a + 23b$$
  
s.t.  $5a + 15b \le 480$   
 $4a + 4b \le 160$   
 $35a + 20b \le 1190$   
 $a, b \ge 0$ 

Note that a lower bound is easy to derive. Every choice of  $a, b \ge 0$  gives us a lower bound (e.g. a = 12, b = 28 gives us a lower bound of 800).

If you take a conic combination of the rows (multiply the i-th row with  $y_i \ge 0$ ) such that  $\sum_i y_i a_{ij} \ge c_j$  then  $\sum_i y_i b_i$  will be an upper bound.

# **Duality**

#### **Definition 2**

Let  $z = \max\{c^T x \mid Ax \le b, x \ge 0\}$  be a linear program P (called the primal linear program).

The linear program D defined by

$$w = \min\{b^T y \mid A^T y \ge c, y \ge 0\}$$

is called the dual problem.

# **Duality**

#### Lemma 3

The dual of the dual problem is the primal problem.

#### **Proof:**

- $w = -\max\{-b^T y \mid -A^T y \le -c, y \ge 0\}$

#### The dual problem is

- $z = -\min\{-c^T x \mid -Ax \ge -b, x \ge 0\}$
- $\triangleright z = \max\{c^T x \mid Ax \le b, x \ge 0\}$

# **Weak Duality**

Let  $z = \max\{c^T x \mid Ax \le b, x \ge 0\}$  and  $w = \min\{b^T y \mid A^T y \ge c, y \ge 0\}$  be a primal dual pair.

x is primal feasible iff  $x \in \{x \mid Ax \le b, x \ge 0\}$ 

y is dual feasible, iff  $y \in \{y \mid A^T y \ge c, y \ge 0\}$ .

#### **Theorem 4 (Weak Duality)**

Let  $\hat{x}$  be primal feasible and let  $\hat{y}$  be dual feasible. Then

$$c^T\hat{x} \leq z \leq w \leq b^T\hat{y} \ .$$

# **Weak Duality**

$$A^T\hat{y} \geq c \Rightarrow \hat{x}^TA^T\hat{y} \geq \hat{x}^Tc \ (\hat{x} \geq 0)$$

$$A\hat{x} \le b \Rightarrow y^T A \hat{x} \le \hat{y}^T b \ (\hat{y} \ge 0)$$

This gives

$$c^T\hat{x} \leq \hat{y}^T A \hat{x} \leq b^T \hat{y} \ .$$

Since, there exists primal feasible  $\hat{x}$  with  $c^T\hat{x}=z$ , and dual feasible  $\hat{y}$  with  $b^T\hat{y}=w$  we get  $z\leq w$ .

If *P* is unbounded then *D* is infeasible.

# 5.2 Simplex and Duality

The following linear programs form a primal dual pair:

$$z = \max\{c^T x \mid Ax = b, x \ge 0\}$$
$$w = \min\{b^T y \mid A^T y \ge c\}$$

This means for computing the dual of a standard form LP, we do not have non-negativity constraints for the dual variables.

### **Proof**

#### Primal:

$$\max\{c^{T}x \mid Ax = b, x \ge 0\}$$

$$= \max\{c^{T}x \mid Ax \le b, -Ax \le -b, x \ge 0\}$$

$$= \max\{c^{T}x \mid \begin{bmatrix} A \\ -A \end{bmatrix} x \le \begin{bmatrix} b \\ -b \end{bmatrix}, x \ge 0\}$$

#### **Dual:**

$$\min\{ \begin{bmatrix} b^T - b^T \end{bmatrix} y \mid \begin{bmatrix} A^T - A^T \end{bmatrix} y \ge c, y \ge 0 \}$$

$$= \min \left\{ \begin{bmatrix} b^T - b^T \end{bmatrix} \cdot \begin{bmatrix} y^+ \\ y^- \end{bmatrix} \mid \begin{bmatrix} A^T - A^T \end{bmatrix} \cdot \begin{bmatrix} y^+ \\ y^- \end{bmatrix} \ge c, y^- \ge 0, y^+ \ge 0 \right\}$$

$$= \min \left\{ b^T \cdot (y^+ - y^-) \mid A^T \cdot (y^+ - y^-) \ge c, y^- \ge 0, y^+ \ge 0 \right\}$$

$$= \min \left\{ b^T y' \mid A^T y' \ge c \right\}$$

# **Proof of Optimality Criterion for Simplex**

Suppose that we have a basic feasible solution with reduced cost

$$\tilde{c} = c^T - c_B^T A_B^{-1} A \le 0$$

This is equivalent to  $A^T(A_R^{-1})^T c_B \ge c$ 

$$y^* = (A_B^{-1})^T c_B$$
 is solution to the dual  $\min\{b^T y | A^T y \ge c\}$ .

$$b^{T}y^{*} = (Ax^{*})^{T}y^{*} = (A_{B}x_{B}^{*})^{T}y^{*}$$
$$= (A_{B}x_{B}^{*})^{T}(A_{B}^{-1})^{T}c_{B} = (x_{B}^{*})^{T}A_{B}^{T}(A_{B}^{-1})^{T}c_{B}$$
$$= c^{T}x^{*}$$

Hence, the solution is optimal.

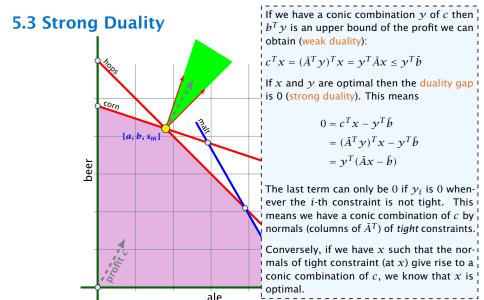
# **5.3 Strong Duality**

$$P = \max\{c^T x \mid Ax \le b, x \ge 0\}$$
  
 $n_A$ : number of variables,  $m_A$ : number of constraints

We can put the non-negativity constraints into A (which gives us unrestricted variables):  $\bar{P} = \max\{c^T x \mid \bar{A}x \leq \bar{b}\}$ 

$$n_{\bar{A}}=n_A$$
,  $m_{\bar{A}}=m_A+n_A$ 

Dual 
$$D = \min\{\bar{b}^T y \mid \bar{A}^T y = c, y \ge 0\}.$$



The profit vector c lies in the cone generated by the normals for the hops and the corn constraint (the tight constraints).

## **Strong Duality**

#### **Theorem 5 (Strong Duality)**

Let P and D be a primal dual pair of linear programs, and let  $z^*$  and  $w^*$  denote the optimal solution to P and D, respectively. Then

$$z^* = w^*$$

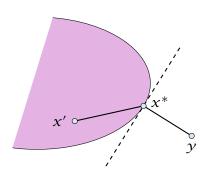
#### Lemma 6 (Weierstrass)

Let X be a compact set and let f(x) be a continuous function on X. Then  $\min\{f(x):x\in X\}$  exists.

(without proof)

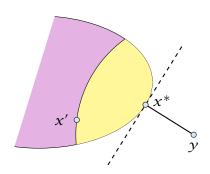
#### Lemma 7 (Projection Lemma)

Let  $X \subseteq \mathbb{R}^m$  be a non-empty convex set, and let  $y \notin X$ . Then there exist  $x^* \in X$  with minimum distance from y. Moreover for all  $x \in X$  we have  $(y - x^*)^T (x - x^*) \le 0$ .



# **Proof of the Projection Lemma**

- ▶ Define f(x) = ||y x||.
- ightharpoonup We want to apply Weierstrass but X may not be bounded.
- ▶  $X \neq \emptyset$ . Hence, there exists  $x' \in X$ .
- ▶ Define  $X' = \{x \in X \mid \|y x\| \le \|y x'\|\}$ . This set is closed and bounded.
- Applying Weierstrass gives the existence.



# **Proof of the Projection Lemma (continued)**

 $x^*$  is minimum. Hence  $||y - x^*||^2 \le ||y - x||^2$  for all  $x \in X$ .

By convexity:  $x \in X$  then  $x^* + \epsilon(x - x^*) \in X$  for all  $0 \le \epsilon \le 1$ .

$$||y - x^*||^2 \le ||y - x^* - \epsilon(x - x^*)||^2$$

$$= ||y - x^*||^2 + \epsilon^2 ||x - x^*||^2 - 2\epsilon(y - x^*)^T (x - x^*)$$

Hence, 
$$(y - x^*)^T (x - x^*) \le \frac{1}{2} \epsilon ||x - x^*||^2$$
.

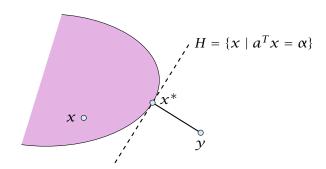
Letting  $\epsilon \to 0$  gives the result.

#### **Theorem 8 (Separating Hyperplane)**

Let  $X \subseteq \mathbb{R}^m$  be a non-empty closed convex set, and let  $y \notin X$ . Then there exists a separating hyperplane  $\{x \in \mathbb{R} : a^Tx = \alpha\}$  where  $a \in \mathbb{R}^m$ ,  $\alpha \in \mathbb{R}$  that separates y from X.  $(a^Ty < \alpha; a^Tx \ge \alpha)$  for all  $x \in X$ 

# **Proof of the Hyperplane Lemma**

- Let  $x^* \in X$  be closest point to y in X.
- ▶ By previous lemma  $(y x^*)^T (x x^*) \le 0$  for all  $x \in X$ .
- Choose  $a = (x^* y)$  and  $\alpha = a^T x^*$ .
- For  $x \in X$ :  $a^T(x x^*) \ge 0$ , and, hence,  $a^Tx \ge \alpha$ .
- Also,  $a^T y = a^T (x^* a) = \alpha ||a||^2 < \alpha$



#### Lemma 9 (Farkas Lemma)

Let A be an  $m \times n$  matrix,  $b \in \mathbb{R}^m$ . Then exactly one of the following statements holds.

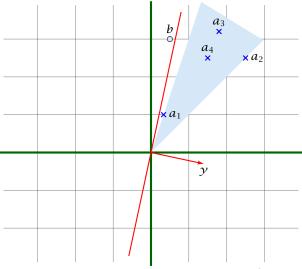
- **1.**  $\exists x \in \mathbb{R}^n$  with Ax = b,  $x \ge 0$
- **2.**  $\exists y \in \mathbb{R}^m$  with  $A^T y \ge 0$ ,  $b^T y < 0$

Assume  $\hat{x}$  satisfies 1. and  $\hat{y}$  satisfies 2. Then

$$0 > y^T b = y^T A x \ge 0$$

Hence, at most one of the statements can hold.

### **Farkas Lemma**



If b is not in the cone generated by the columns of A, there exists a hyperplane y that separates b from the cone.

### **Proof of Farkas Lemma**

Now, assume that 1. does not hold.

Consider  $S = \{Ax : x \ge 0\}$  so that S closed, convex,  $b \notin S$ .

We want to show that there is y with  $A^Ty \ge 0$ ,  $b^Ty < 0$ .

Let y be a hyperplane that separates b from S. Hence,  $y^Tb < \alpha$  and  $y^Ts \ge \alpha$  for all  $s \in S$ .

$$0 \in S \Rightarrow \alpha \le 0 \Rightarrow y^T b < 0$$

 $y^T A x \ge \alpha$  for all  $x \ge 0$ . Hence,  $y^T A \ge 0$  as we can choose x arbitrarily large.

#### Lemma 10 (Farkas Lemma; different version)

Let A be an  $m \times n$  matrix,  $b \in \mathbb{R}^m$ . Then exactly one of the following statements holds.

- **1.**  $\exists x \in \mathbb{R}^n$  with  $Ax \leq b$ ,  $x \geq 0$
- **2.**  $\exists y \in \mathbb{R}^m$  with  $A^T y \ge 0$ ,  $b^T y < 0$ ,  $y \ge 0$

#### Rewrite the conditions:

1. 
$$\exists x \in \mathbb{R}^n \text{ with } [A \ I] \cdot \begin{vmatrix} x \\ s \end{vmatrix} = b, \ x \ge 0, \ s \ge 0$$

**2.** 
$$\exists y \in \mathbb{R}^m \text{ with } \begin{bmatrix} A^T \\ I \end{bmatrix} y \ge 0, b^T y < 0$$

$$P: z = \max\{c^T x \mid Ax \le b, x \ge 0\}$$

*D*: 
$$w = \min\{b^T y \mid A^T y \ge c, y \ge 0\}$$

### **Theorem 11 (Strong Duality)**

Let P and D be a primal dual pair of linear programs, and let z and w denote the optimal solution to P and D, respectively (i.e., P and D are non-empty). Then

$$z = w$$
.

 $z \leq w$ : follows from weak duality

 $z \geq w$ :

We show  $z < \alpha$  implies  $w < \alpha$ .

$$\exists x \in \mathbb{R}^n$$
s.t. 
$$Ax \leq b$$

$$-c^T x \leq -\alpha$$

$$x \geq 0$$

$$\exists y \in \mathbb{R}^m; v \in \mathbb{R}$$
s.t.  $A^T y - cv \ge 0$ 

$$b^T y - \alpha v < 0$$

$$y, v \ge 0$$

From the definition of  $\alpha$  we know that the first system is infeasible; hence the second must be feasible.

$$\exists y \in \mathbb{R}^m; v \in \mathbb{R}$$
s.t.  $A^T y - cv \ge 0$ 

$$b^T y - \alpha v < 0$$

$$y, v \ge 0$$

If the solution y, v has v = 0 we have that

$$\exists y \in \mathbb{R}^m$$
s.t.  $A^T y \ge 0$ 

$$b^T y < 0$$

$$y \ge 0$$

is feasible. By Farkas lemma this gives that LP  ${\it P}$  is infeasible. Contradiction to the assumption of the lemma.

Hence, there exists a solution y, v with v > 0.

We can rescale this solution (scaling both y and v) s.t. v = 1.

Then y is feasible for the dual but  $b^Ty < \alpha$ . This means that  $w < \alpha$ .

### **Fundamental Questions**

#### **Definition 12 (Linear Programming Problem (LP))**

Let  $A \in \mathbb{Q}^{m \times n}$ ,  $b \in \mathbb{Q}^m$ ,  $c \in \mathbb{Q}^n$ ,  $\alpha \in \mathbb{Q}$ . Does there exist  $x \in \mathbb{Q}^n$  s.t. Ax = b,  $x \ge 0$ ,  $c^T x \ge \alpha$ ?

#### Questions:

- ► Is LP in NP?
- Is LP in co-NP? yes!
- ▶ Is LP in P?

#### Proof:

- Given a primal maximization problem P and a parameter  $\alpha$ . Suppose that  $\alpha > \text{opt}(P)$ .
- We can prove this by providing an optimal basis for the dual.
- A verifier can check that the associated dual solution fulfills all dual constraints and that it has dual cost  $< \alpha$ .



## **Complementary Slackness**

#### Lemma 13

Assume a linear program  $P = \max\{c^Tx \mid Ax \leq b; x \geq 0\}$  has solution  $x^*$  and its dual  $D = \min\{b^Ty \mid A^Ty \geq c; y \geq 0\}$  has solution  $y^*$ .

- **1.** If  $x_j^* > 0$  then the *j*-th constraint in *D* is tight.
- **2.** If the *j*-th constraint in *D* is not tight than  $x_i^* = 0$ .
- **3.** If  $y_i^* > 0$  then the *i*-th constraint in *P* is tight.
- **4.** If the *i*-th constraint in *P* is not tight than  $y_i^* = 0$ .

If we say that a variable  $x_j^*$  ( $y_i^*$ ) has slack if  $x_j^* > 0$  ( $y_i^* > 0$ ), (i.e., the corresponding variable restriction is not tight) and a contraint has slack if it is not tight, then the above says that for a primal-dual solution pair it is not possible that a constraint **and** its corresponding (dual) variable has slack.

# **Proof: Complementary Slackness**

Analogous to the proof of weak duality we obtain

$$c^T x^* \le y^{*T} A x^* \le b^T y^*$$

Because of strong duality we then get

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

This gives e.g.

$$\sum_{j} (y^T A - c^T)_j x_j^* = 0$$

From the constraint of the dual it follows that  $y^TA \ge c^T$ . Hence the left hand side is a sum over the product of non-negative numbers. Hence, if e.g.  $(y^TA - c^T)_j > 0$  (the j-th constraint in the dual is not tight) then  $x_j = 0$  (2.). The result for (1./3./4.) follows similarly.

## Interpretation of Dual Variables

Brewer: find mix of ale and beer that maximizes profits

max 
$$13a + 23b$$
  
s.t.  $5a + 15b \le 480$   
 $4a + 4b \le 160$   
 $35a + 20b \le 1190$   
 $a, b \ge 0$ 

► Entrepeneur: buy resources from brewer at minimum cost *C*, *H*, *M*: unit price for corn, hops and malt.

min 
$$480C$$
 +  $160H$  +  $1190M$   
s.t.  $5C$  +  $4H$  +  $35M \ge 13$   
 $15C$  +  $4H$  +  $20M \ge 23$   
 $C, H, M \ge 0$ 

Note that brewer won't sell (at least not all) if e.g. 5C+4H+35M<13 as then brewing ale would be advantageous.

## Interpretation of Dual Variables

#### **Marginal Price:**

- How much money is the brewer willing to pay for additional amount of Corn, Hops, or Malt?
- ▶ We are interested in the marginal price, i.e., what happens if we increase the amount of Corn, Hops, and Malt by  $\varepsilon_C$ ,  $\varepsilon_H$ , and  $\varepsilon_M$ , respectively.

The profit increases to  $\max\{c^Tx\mid Ax\leq b+\varepsilon; x\geq 0\}$ . Because of strong duality this is equal to

min 
$$(b^T + \epsilon^T)y$$
  
s.t.  $A^Ty \ge c$   
 $y \ge 0$ 

## **Interpretation of Dual Variables**

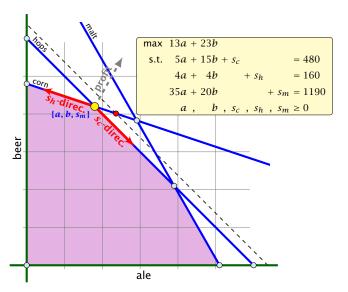
If  $\epsilon$  is "small" enough then the optimum dual solution  $y^*$  might not change. Therefore the profit increases by  $\sum_i \epsilon_i y_i^*$ .

Therefore we can interpret the dual variables as marginal prices.

Note that with this interpretation, complementary slackness becomes obvious.

- ▶ If the brewer has slack of some resource (e.g. corn) then he is not willing to pay anything for it (corresponding dual variable is zero).
- ▶ If the dual variable for some resource is non-zero, then an increase of this resource increases the profit of the brewer. Hence, it makes no sense to have left-overs of this resource. Therefore its slack must be zero.

## **Example**



The change in profit when increasing hops by one unit is

$$=\underbrace{c_B^T A_B^{-1}}_{v^*} e_h.$$

Of course, the previous argument about the increase in the primal objective only holds for the non-degenerate case.

If the optimum basis is degenerate then increasing the supply of one resource may not allow the objective value to increase.

### **Flows**

#### **Definition 14**

An (s,t)-flow in a (complete) directed graph  $G=(V,V\times V,c)$  is a function  $f:V\times V\mapsto \mathbb{R}^+_0$  that satisfies

**1.** For each edge (x, y)

$$0 \le f_{xy} \le c_{xy} .$$

(capacity constraints)

**2.** For each  $v \in V \setminus \{s, t\}$ 

$$\sum_{x} f_{vx} = \sum_{x} f_{xv} .$$

(flow conservation constraints)



### **Flows**

#### **Definition 15**

The value of an (s, t)-flow f is defined as

$$\operatorname{val}(f) = \sum_{x} f_{sx} - \sum_{x} f_{xs}$$
.

#### **Maximum Flow Problem:**

Find an (s, t)-flow with maximum value.

$$\begin{array}{llll} & & \sum_{(xy)} c_{xy} \ell_{xy} \\ \text{s.t.} & f_{xy} \left( x, y \neq s, t \right) \colon & 1 \ell_{xy} - 1 p_x + 1 p_y \; \geq \; 0 \\ & f_{sy} \left( y \neq s, t \right) \colon & 1 \ell_{sy} \; + 1 p_y \; \geq \; 1 \\ & f_{xs} \left( x \neq s, t \right) \colon & 1 \ell_{xs} - 1 p_x \; \; \geq \; -1 \\ & f_{ty} \left( y \neq s, t \right) \colon & 1 \ell_{ty} \; + 1 p_y \; \geq \; 0 \\ & f_{xt} \left( x \neq s, t \right) \colon & 1 \ell_{xt} - 1 p_x \; \; \geq \; 0 \\ & f_{st} \colon & 1 \ell_{st} \; \; \geq \; 1 \\ & f_{ts} \colon & 1 \ell_{ts} \; \; \geq \; -1 \\ & \ell_{xy} \; \; \geq \; 0 \end{array}$$

min		$\sum_{(xy)} c_{xy} \ell_{xy}$	
s.t.	$f_{xy}(x, y \neq s, t)$ :	$1\ell_{xy} - 1p_x + 1p_y \ge 0$	
	$f_{sy} (y \neq s, t)$ :	$1\ell_{sy} - 1 + 1p_y \ge 0$	
	$f_{xs}(x \neq s,t)$ :	$1\ell_{xs} - 1p_x + 1 \ge 0$	
	$f_{ty} (y \neq s, t)$ :	$1\ell_{ty} - 0 + 1p_y \ge 0$	
	$f_{xt}$ $(x \neq s, t)$ :	$1\ell_{xt} - 1p_x + 0 \ge 0$	
	$f_{st}$ :	$1\ell_{st}$ $-1+0 \geq 0$	
	$f_{ts}$ :	$1\ell_{ts}$ $ 0+$ $1 \geq 0$	
		$\ell_{xy} \geq 0$	

$$\begin{array}{llll} & & \sum_{(xy)} c_{xy} \ell_{xy} \\ & \text{s.t.} & f_{xy} \; (x,y \neq s,t) \colon & 1\ell_{xy} - 1p_x + 1p_y \; \geq \; 0 \\ & f_{sy} \; (y \neq s,t) \colon & 1\ell_{sy} - \; p_s + 1p_y \; \geq \; 0 \\ & f_{xs} \; (x \neq s,t) \colon & 1\ell_{xs} - 1p_x + \; p_s \; \geq \; 0 \\ & f_{ty} \; (y \neq s,t) \colon & 1\ell_{ty} - \; p_t + 1p_y \; \geq \; 0 \\ & f_{xt} \; (x \neq s,t) \colon & 1\ell_{xt} - 1p_x + \; p_t \; \geq \; 0 \\ & f_{st} \colon & 1\ell_{st} - \; p_s + \; p_t \; \geq \; 0 \\ & f_{ts} \colon & 1\ell_{ts} - \; p_t + \; p_s \; \geq \; 0 \\ & \ell_{xy} \; \geq \; 0 \end{array}$$

with  $p_t = 0$  and  $p_s = 1$ .

min 
$$\sum_{(xy)} c_{xy} \ell_{xy}$$
s.t.  $f_{xy}$ :  $1\ell_{xy} - 1p_x + 1p_y \ge 0$ 

$$\ell_{xy} \ge 0$$

$$p_s = 1$$

$$p_t = 0$$

We can interpret the  $\ell_{xy}$  value as assigning a length to every edge.

The value  $p_x$  for a variable, then can be seen as the distance of x to t (where the distance from s to t is required to be 1 since  $p_s = 1$ ).

The constraint  $p_x \le \ell_{xy} + p_y$  then simply follows from triangle inequality  $(d(x,t) \le d(x,y) + d(y,t) \Rightarrow d(x,t) \le \ell_{xy} + d(y,t))$ .



One can show that there is an optimum LP-solution for the dual problem that gives an integral assignment of variables.

This means  $p_{\chi}=1$  or  $p_{\chi}=0$  for our case. This gives rise to a cut in the graph with vertices having value 1 on one side and the other vertices on the other side. The objective function then evaluates the capacity of this cut.

This shows that the Maxflow/Mincut theorem follows from linear programming duality.