LP. Lecture 3. Chapter 3: degeneracy.

- degeneracy
- example cycling
- the lexicographic method
- other pivot rules
- the fundamental theorem of LP

Repetition

- the simplex algorithm: sequence of pivots starting with initial feasible basic solution
- the simplex method: 2 times the simplex algorithm: Phase I and Phase II
- Phase I: solve the Phase I problem to find, if possible, an initial feasible solution. In that case we also have a basic feasible solution for the next task.
- Phase II: solve Phase II problem with the solution from the Phase I problem as a starting point. Using the simplex algorithm we then find an optimal solution of the original problem or ....
- unbounded problem: ... no basic variable goes towards zero, and we find a ray along which the objective function η goes towards infinity.

Degeneracy

Consider the LP problem:

$$\begin{array}{lrll} \text{max} & \sum_{j=1}^{n} c_j x_j & & \\ \text{subject to} & & & \\ & \sum_{j=1}^{n} a_{i,j} x_j & \leq b_i & \text{for } i = 1, \dots, m \\ & x_j & \geq 0 & \text{for } j = 1, \dots, n. \end{array}$$

By the start of a pivot we have the dictionary:

$$\begin{array}{rcl} \eta & = & \bar{\eta} + \sum_{j \in N} \bar{c}_j x_j \\ x_i & = & \bar{b}_i - \sum_{j \in N} \bar{a}_{i,j} x_j \quad \text{for } i \in B. \end{array}$$

We call the dictionary degenerate if $\bar{b}_i = 0$ for at least one i. We then have a degenerate basic solution. Note that we always have $\bar{b}_i \geq 0$. *Why?*

The reason is: we start with each $\bar{b}_i \geq 0$ and this property is maintained during the iterations since we do not increase the entering variable too much. So far we have assumed that $\bar{b}_i > 0$.

Consequence: we have always been able to increase the entering variable by a *positive amount*.

Degeneracy and problems:

- degeneracy may not cause problems
- degeneracy could give problems.

For instance a degenerate dictionary may be optimal:

$$\begin{array}{rcrcrcr} \eta & = & 2 & - & 3x_4 & - & x_5 \\ \hline x_1 & = & 1 & + & x_4 & + & x_5 \\ x_2 & = & 2 & + & 5x_4 & + & x_5 \\ x_3 & = & 0 & & & - & x_5 \end{array}$$

No problem with this! And a similar remark holds for unbounded problem.

But: if a degenerate dictionary gives a degenerate pivot, we may get into trouble. This means that the entering variable cannot be increased, $\theta = 0$. Example:

$$\begin{array}{rcrcrcr}
\eta & = & 2 & + & 3x_4 & - & x_5 \\
\hline
x_1 & = & 0 & - & x_4 & + & x_5 \\
x_2 & = & 2 & + & 5x_4 & + & x_5 \\
x_3 & = & 0 & & & - & x_5
\end{array}$$

- We want to increase x_4, but this cannot be done as x_1 then becomes negative.
- However, we can still make this degenerate pivot by taking x_4 into the basis and x_1 out of the basis. But the *solution* x is still the same and therefore η is unchanged.
- So: we have done a pivot, but we have the same point $x \in \mathbb{R}^n$. Geometrically, there is no change, but algebraically there is: we have a new basis.

Do degenerate pivots cause problems?

Note necessarily. But if we have a number of denerate pivots and return to the *same* basis, then this cycle of pivots would be repeated infinitely many times. So the algorithm would get trapped. This phenomenon is called cycling.

Comments on degeneracy and cycling:

- cycling is hardly any problem in practice. For LP problems in practice one has hardly ever seen that cycling occuring.
- degeneracy arises frequently in practical LP problems. So, several or even most of the pivots may be degenerate. But this feature seems to be hard to avoid (after several attempts duringf the development of LP).
- cycling may happen. The first such example was constructed in 1953 by Alan Hoffman.

Cycling, example

Let us use as pivot rule: choose as entering variable the nonbasic variable with $\bar{c}_j$ largest possible, and as leaving variable the one with the smallest index.

Dictionary 0:

$$\begin{array}{rcrcrcrcr}
\eta & = 0 & & 10x_1 & - & 57x_2 & - & 9x_3 & - & 24x_4 \\
\hline
w_1 & = 0 & - & 0.5x_1 & + & 5.5x_2 & + & 2.5x_3 & - & 9x_4 \\
w_2 & = 0 & - & 0.5x_1 & + & 1.5x_2 & + & 0.5x_3 & - & x_4 \\
w_3 & = 1 & - & x_1 & & & & & &
\end{array}$$

Dictionary 1:

$$\begin{array}{rcrcrcrcr}
\eta & = 0 & - & 20w_1 & + & 53x_2 & + & 41x_3 & - & 204x_4 \\
\hline
x_1 & = 0 & - & 2w_1 & + & 11x_2 & + & 5x_3 & - & 18x_4 \\
w_2 & = 0 & + & w_1 & - & 4x_2 & - & 2x_3 & + & 8x_4 \\
w_3 & = 1 & + & 2w_1 & - & 11x_2 & - & 5x_3 & + & 18x_4
\end{array}$$

Dictionary 2:

$$\begin{array}{lcrcrcrcr}
\eta & = 0 & - & 6.75w_1 & - & 13.25w_2 & + & 14.5x_3 & - & 98x_4 \\
\hline
x_1 & = 0 & + & 0.75w_1 & - & 2.75w_2 & - & 0.5x_3 & + & 4x_4 \\
x_2 & = 0 & + & 0.25w_1 & - & 0.254w_2 & - & 0.5x_3 & + & 2x_4 \\
w_3 & = 1 & - & 0.75w_1 & - & 13.25w_2 & + & 0.5x_3 & - & 4x_4
\end{array}$$

Dictionary 3:

$$\begin{array}{lcrcrcrcr}
\eta & = 0 & + & 15w_1 & - & 93w_2 & - & 29x_1 & + & 18x_4 \\
\hline
x_3 & = 0 & + & 1.5w_1 & - & 5.5w_2 & - & 2x_1 & + & 8x_4 \\
x_2 & = 0 & - & 0.5w_1 & + & 2.5w_2 & + & x_1 & - & 2x_4 \\
w_3 & = 1 & & & & & - & x_1 & &
\end{array}$$

Dictionary 4:

$$\begin{array}{lllllllll}
\eta & = 0 & + & 10.5w_1 & - & 70.5w_2 & - & 20x_1 & - & 9x_2 \\
\hline
x_3 & = 0 & - & 0.5w_1 & + & 4.5w_2 & + & 2x_1 & - & 4x_2 \\
x_4 & = 0 & - & 0.25w_1 & + & 1.25w_2 & + & 0.5x_1 & - & 0.5x_2 \\
w_3 & = 1 & & & & & - & x_1 & &
\end{array}$$

Dictionary 5:

$$\begin{array}{lllllllll}
\eta & = 0 & - & 21x_3 & + & 24w_2 & + & 22x_1 & - & 93x_2 \\
\hline
w_1 & = 0 & - & 2x_3 & + & 9w_2 & + & 4x_1 & - & 8x_2 \\
x_4 & = 0 & + & 0.5x_3 & - & w_2 & - & 0.5x_1 & + & 1.5x_2 \\
w_3 & = 1 & & & & & - & x_1 & &
\end{array}$$

Dictionary 6:

$$
\begin{array}{lllllllll}
\eta & = 0 & + & 10x_1 & - & 57x_2 & - & 9x_3 & - & 24x_4 \\
\hline
w_1 & = 0 & - & 0.5x_1 & + & 5.5x_2 & + & 2.5x_3 & - & 9x_4 \\
w_2 & = 0 & - & 0.5x_1 & + & 1.5x_2 & + & 0.5x_3 & - & x_4 \\
w_3 & = 1 & - & x_1 & & & & & &
\end{array}
$$

We see that dictionary 6 is the same as dictionary 0. So we have cycling!

Cycling is therefore a (theoretical) problem. However, cycling is in fact the *only* problem that arises because we have the following result ....

An important theorem

Theorem 3.1. *If the simplex method does not terminate, then it must cycle.*

Proof: How many dictionaries are there? An upper bound is

$$\binom{m+n}{n} = (m+n)!/(n!m!)$$

which is the number of ways of selecting m elements (the basic variables) from $n+m$ elements (all variables). (Remark: we here use that the dictionary is determined by the basic variables (when the equations are ordered e.g. according to the index of the basic variables): this is most easy to verify when we have introduced the matrix version of LP). If the simplex algorithm does not terminate, then two of these dictionaries must occur twice, so we have cycling. □

The lexicographic method

The lexicographic method (or the closely related perturbation method) is a method to avoid cycling.

- The idea: perturb the right-hand sides to avoid degeneracy!
- If these perturbations are small enough, the problem will change so little that we still obtain a correct optimal solution.

Example (with a degenerate basic solution):

$$
\begin{array}{rlrcr}
\eta & = 4 & +2x_1 & - & x_2 \\
\hline
w_1 & = 0.5 & & - & x_2 \\
w_2 & = 0 & -2x_1 & + & 4x_2 \\
w_3 & = 0 & x_1 & - & 3x_2
\end{array}
$$

We now introduce symbols (unspecified small numbers) where

$$0 < \epsilon_3 \ll \epsilon_2 \ll \epsilon_1 \ll \text{all data}$$

and the perturbed dictionary

$$
\begin{array}{lllllrcr}
\eta & = 4 & & & & +2x_1 & - & x_2 \\
\hline
w_1 & = 0.5 & +\epsilon_1 & & & & - & x_2 \\
w_2 & = 0 & & +\epsilon_2 & & -2x_1 & + & 4x_2 \\
w_3 & = 0 & & & +\epsilon_3 & +x_1 & - & 3x_2
\end{array}
$$

Not degenerate! Pivot: x_1 in, and w_2 out.

Next:

$$
\begin{array}{llllll}
\eta & = 4 & & +\epsilon_2 & & +2w_2 + 3x_2 \\
\hline
w_1 & = 0.5 & +\epsilon_1 & & & - x_2 \\
x_1 & = & & 0.5\epsilon_2 & & -0.5w_2 + 2x_2 \\
w_3 & = & & 0.5\epsilon_2 & +\epsilon_3 & -0.5w_2 - x_2
\end{array}
$$

and ...

$$
\begin{array}{lllllll}
\eta & = 4 & & +2.5\epsilon_2 & +3\epsilon_3 & -2.5w_2 & -3w_3 \\
\hline
w_1 & = 0.5 & +\epsilon_1 & -0.5\epsilon_2 & -\epsilon_3 & +0.5w_2 & +w_3 \\
x_1 & = & & 1.5\epsilon_2 & +2\epsilon_3 & -1.5w_2 & -2w_3 \\
x_2 & = & & 0.5\epsilon_2 & +\epsilon_3 & -0.5w_2 & -w_3
\end{array}
$$

This dictionary is optimal.

Now drop the perturbations to obtain

$$
\begin{array}{llll}
\eta & = 4 & -2.5w_2 & -3w_3 \\
\hline
w_1 & = 0.5 & +0.5w_2 & +w_3 \\
x_1 & = 0 & -1.5w_2 & -2w_3 \\
x_2 & = 0 & -0.5w_2 & -w_3
\end{array}
$$

which gives an optimal solution of the original LP problem!

- The perturbations will only affect the constant terms, not the coefficients of the variables. Why?
- Therefore the choice of entering variable is not affected. But the leaving variable is uniquely determined in every iteration. (see below).

Another important theorem!

Theorem 3.2. *The simplex method will always terminate whenever the leaving variable is selected using the lexicographic method.*

Proof: It is enough to show that we never get a degenerate dictionary. Consider the "constant part" which initially is:

$$\begin{matrix} \epsilon_1 & & \\ & \ddots & \\ & & \epsilon_m \end{matrix}$$

or $I\epsilon$ in matrix form (where I is the identity matrix, and ϵ is the column vector with components being the ϵ_i's). During the pivots a multiple of one row is added to other rows, and this corresponds to multiplication from the left by certain nonsingular (i.e., invertible, in fact, elementary) matrices. □

By the start of an arbitrary pivot the constant part is

$$\begin{matrix} r_{11}\epsilon_1 & \dots & r_{im}\epsilon_m \\ \vdots & & \vdots \\ r_{m1}\epsilon_1 & \dots & r_{mm}\epsilon_m \end{matrix}$$

i.e., in matrix form $R\epsilon$. Since R is nonsingular, it has no row equal to the zero vector. Thus we do not get a degenerate dictionary. □

Bland's rule

In 1977 R. Bland published a new and simple pivot rule for the simplex algorithm. Consider an LP problem with variables $(x_1, x_2, \ldots, x_n)$ (where some may be slack variables, this plays no rule here).

Bland's rule:

- If there are several candidates for *entering variable*, always choose the one with smallest index (subscript).
- If there are several candidates for *leaving variable*, always choose the one with smallest index.

For instance, in an iteration where $\bar{c}_3, \bar{c}_5, \bar{c}_9$ are positive, while $\bar{c}_j \leq 0$ for $j \neq 3, 5, 9$, Bland's rule tells us to choose x_3 as the entering variable.
Note: Bland's rule is applied to *both entering and leaving variables*. This is in contrast to the lexicographic rule which only involved a certain choice for the leaving variable.

Theorem 3.3. *The simplex method will always terminate whenever the entering and leaving variable are selected using the Bland's rule.*

The proof is found in Vanderbei's book (it is rather cryptical, but all known proof for this result are, unfortunately.)

Some comments on pivot rules:

Bland's rule:

- Strength: avoids cycling, easy to understand, easy to implement.
- Weakness: one typically gets a small increase in the objective function in each iteration, so many pivots and long computational time.

Dantzig's rule =largest coefficient rule (1951):

- Choose entering variable with $\bar{c}_j$ largest possible, and leaving variable with smallest index.
- This rule is often used in practice, and it works well (small computational time).
- But it might, in theory, give cycling.

There are other pivot rules, e.g.:

The steepest edge rule:

- Choose entering variable such that the angle between c and the direction vector $x^1 - x^0$ is smallest possible. (Why is this done?) Here x^0 and x^1 are old and new basic solution.
- This rule seems to be one of the best ones in practice (but gives no guarantee against cycling).

Best improvement rule:

- Choose entering variable such that the improvement of the objective function is largest possible.
- May seem clever, but takes too much time in each iteration.

The fundamental theorem of LP

There are two theorems in LP that are more important than all others; they are

- The fundamental theorem of LP, and
- The duality theorem.

We are ready for the first of these! (And duality is coming later.)

Theorem 3.3. *For every LP problem the following is true:*

- *If there is no optimal solution, then the problem is either nonfeasible or unbounded.*
- *If the problem is feasible, there exist a basic feasible solution.*
- *If the problem has an optimal solution, then there exist an optimal basic solution.*

Proof: The Phase I problem determines if the original problem is feasible, and if it is, it also finds a basic feasible solution. Then Phase II determines if there is an optimal solution or if the problem is unbounded. These are the only possibilities since the method terminates due to the existence of anticycling rules. □

Geometry and degeneracy

- example: an Egyption pyramid!
- the slack variables "measure" distance to the hyperplane
- degeneracy type 1: redundant inequalities
- degeneracy type 2: $P \subset \mathbb{R}^n$, a vertex is on more than n facets
- degenerate pivots: represents the same vertex x in different ways (there may be several selctions of n linearly independent hyperplanes through x)