FOUNDATIONS OF OPTIMIZATION: IE6001

Degeneracy

Napat Rujeerapaiboon

Semester I, AY2022/2023

Degeneracy

Definition: A BS is called degenerate if more than $n - m$ of its components are zero.

Lemma: Assume that $\forall i = 1, \ldots, n \; \exists \; \hat{x}$ with $A\hat{x} = b$ and $\hat{x}_i \neq 0$. Then, any BS x is degenerate iff it is associated with more than one index set.

Proof: "$\Leftarrow$"

- Suppose a BS x corresponds to I_1 and I_2, $I_1 \neq I_2$.
- Then $x_i = 0$ if either $i \notin I_1$ or $i \notin I_2$.
- As $I_1 \neq I_2$, x has more than $n - m$ zero components.

$\Rightarrow$ x is a degenerate BS.

Proof: "$\Rightarrow$"

- Suppose x is a degenerate BS associated with some index set I; consider the corresponding simplex tableau.
- By degeneracy, $\exists p \in I$ with $y_{p0} = 0$.
- $\exists q \notin I$ such that $y_{pq} \neq 0$. Otherwise, $x_p = 0$ is part of the representation, but this is not possible by assumption.
- Pivoting on (p, q) gives a new basic solution which is identical to current one since

$$y'_{q0} = \frac{y_{p0}}{y_{pq}} = y_{p0} = 0 \text{ and } y'_{i0} = y_{i0} - \frac{y_{iq}}{y_{pq}} y_{p0} = y_{i0} \ \forall i \in I \backslash \{p\}.$$

$\Rightarrow$ x corresponds to the index sets I and $(I \backslash \{p\}) \cup \{q\}$.

$\square$

Degeneracy (cont)

Note: The index sets I and $(I \backslash \{p\}) \cup \{q\}$ produce the same BS but different basic representations.

Degeneracy and Simplex Algorithm

How does degeneracy affect the simplex algorithm?

- If we pivot on (p, q) and if $y_{p0} = 0$, then the new BS is identical to old one.
- In particular, we find

$$\beta_0' = \beta_0 - \frac{\beta_q}{y_{pq}} y_{p0} = \beta_0,$$

  and the finite termination theorem breaks down (no strict improvement of objective value).
- A pivot step (p, q) is called degenerate if $y_{p0} = 0$ and non-degenerate otherwise.

Degeneracy and Simplex Algorithm

The simplex algorithm can now be decomposed into:

$$\left[\begin{array}{c}\text{sequence of}\\ \text{degenerate}\\ \text{pivots}\end{array}\right] \quad \begin{array}{c}\text{non-}\\ \text{degenerate}\\ \text{pivot}\end{array} \quad \left[\begin{array}{c}\text{sequence of}\\ \text{degenerate}\\ \text{pivots}\end{array}\right] \ldots$$

Note: Some or all of these sequences of degenerate pivots may be empty.

Geometrically, the current BFS remains unchanged throughout a sequence of degenerate pivots, and a non-degenerate pivot moves it to a different BFS.

Degeneracy and Simplex Algorithm

- We know that the number of index sets is $\leq \binom{n}{m}$.

$\Rightarrow$ Sequences of degenerate pivots are finite if no index set is repeated.

- However, suppose that the index sets $I_1, I_2, \ldots, I_k, \ldots$ correspond to a sequence of degenerate pivots and that $I_k = I_{k+\ell}$ for some $\ell \in \mathbb{N}$.
- If a given index set determines a unique pivot, we have

$$\begin{array}{ccccccc} I_k & = & I_{k+\ell} & = & I_{k+2\ell} & = & \ldots \\ I_{k+1} & = & I_{k+\ell+1} & = & I_{k+2\ell+1} & = & \ldots \\ I_{k+2} & = & I_{k+\ell+2} & = & I_{k+2\ell+2} & = & \ldots \end{array}$$

and so the algorithm will cycle and never terminate!

Standard Pivoting Conventions (rep)

- If there are several $q \notin I$ with $\beta_q > 0$, then choose q with

$$\beta_q = \max_{j \notin I} \left\{ \beta_j \right\} . \qquad (*)$$

  Such a q produces the maximum decrease in x_0 per unit of increase in x_q.
- If several $q \notin I$ satisfy $(*)$, choose the one with the smallest index.
- If there are several $p \in \arg\min_{i \in I} \overline{x}_{iq}$, choose the one with the smallest index.

Example

$$\min x_0 = -\frac{3}{4}x_4 + 20x_5 - \frac{1}{2}x_6 + 6x_7$$

subject to:

$$\begin{array}{llllllll}
x_1 & & & +\frac{1}{4}x_4 & -8x_5 & -x_6 & +9x_7 & = 0 \\
& x_2 & & +\frac{1}{2}x_4 & -12x_5 & -\frac{1}{2}x_6 & +3x_7 & = 0 \\
& & x_3 & & & +x_6 & & = 1
\end{array}$$

Use standard pivoting conventions!

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	0	0	0	$\frac{3}{4}$	-20	$\frac{1}{2}$	-6	0
x_1	1	0	0	$\frac{1}{4}$	-8	-1	9	0
x_2	0	1	0	$\frac{1}{2}$	-12	$-\frac{1}{2}$	3	0
x_3	0	0	1	0	0	1	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	0	0	0	$\frac{3}{4}$	-20	$\frac{1}{2}$	-6	0
x_1	1	0	0	$\frac{1}{4}$	-8	-1	9	0
x_2	0	1	0	$\frac{1}{2}$	-12	$-\frac{1}{2}$	3	0
x_3	0	0	1	0	0	1	0	1
x_0	-3	0	0	0	4	$\frac{7}{2}$	-33	0
x_4	4	0	0	1	-32	-4	36	0
x_2	-2	1	0	0	4	$\frac{3}{2}$	-15	0
x_3	0	0	1	0	0	1	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	-3	0	0	0	4	$\frac{7}{2}$	-33	0
x_4	4	0	0	1	-32	-4	36	0
x_2	-2	1	0	0	4	$\frac{3}{2}$	-15	0
x_3	0	0	1	0	0	1	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	-3	0	0	0	4	$\frac{7}{2}$	-33	0
x_4	4	0	0	1	-32	-4	36	0
x_2	-2	1	0	0	4	$\frac{3}{2}$	-15	0
x_3	0	0	1	0	0	1	0	1
x_0	-1	-1	0	0	0	2	-18	0
x_4	-12	8	0	1	0	8	-84	0
x_5	$-\frac{1}{2}$	$\frac{1}{4}$	0	0	1	$\frac{3}{8}$	$-\frac{15}{4}$	0
x_3	0	0	1	0	0	1	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	-1	-1	0	0	0	2	-18	0
x_4	-12	8	0	1	0	8	-84	0
x_5	$-\frac{1}{2}$	$\frac{1}{4}$	0	0	1	$\frac{3}{8}$	$-\frac{15}{4}$	0
x_3	0	0	1	0	0	1	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	-1	-1	0	0	0	2	-18	0
x_4	-12	8	0	1	0	8	-84	0
x_5	$-\frac{1}{2}$	$\frac{1}{4}$	0	0	1	$\frac{3}{8}$	$-\frac{15}{4}$	0
x_3	0	0	1	0	0	1	0	1
x_0	2	-3	0	$-\frac{1}{4}$	0	0	3	0
x_6	$-\frac{3}{2}$	1	0	$\frac{1}{8}$	0	1	$-\frac{21}{2}$	0
x_5	$\frac{1}{16}$	$-\frac{1}{8}$	0	$\frac{3}{64}$	1	0	$\frac{3}{16}$	0
x_3	$\frac{3}{2}$	-1	1	$-\frac{1}{8}$	0	0	$\frac{21}{2}$	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	2	-3	0	$-\frac{1}{4}$	0	0	3	0
x_6	$-\frac{3}{2}$	1	0	$\frac{1}{8}$	0	1	$-\frac{21}{2}$	0
x_5	$\frac{1}{16}$	$-\frac{1}{8}$	0	$\frac{3}{64}$	1	0	$\frac{3}{16}$	0
x_3	$\frac{3}{2}$	-1	1	$-\frac{1}{8}$	0	0	$\frac{21}{2}$	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	2	-3	0	$-\frac{1}{4}$	0	0	3	0
x_6	$-\frac{3}{2}$	1	0	$\frac{1}{8}$	0	1	$-\frac{21}{2}$	0
x_5	$\frac{1}{16}$	$-\frac{1}{8}$	0	$\frac{3}{64}$	1	0	$\frac{3}{16}$	0
x_3	$\frac{3}{2}$	-1	1	$-\frac{1}{8}$	0	0	$\frac{21}{2}$	1
x_0	1	-1	0	$\frac{1}{2}$	-16	0	0	0
x_6	2	-6	0	$-\frac{5}{2}$	56	1	0	0
x_7	$\frac{1}{3}$	$-\frac{2}{3}$	0	$-\frac{1}{4}$	$\frac{16}{3}$	0	1	0
x_3	-2	6	1	$\frac{5}{2}$	-56	0	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	1	-1	0	$\frac{1}{2}$	-16	0	0	0
x_6	2	-6	0	$-\frac{5}{2}$	56	1	0	0
x_7	$\frac{1}{3}$	$-\frac{2}{3}$	0	$-\frac{1}{4}$	$\frac{16}{3}$	0	1	0
x_3	-2	6	1	$\frac{5}{2}$	-56	0	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	1	-1	0	$\frac{1}{2}$	-16	0	0	0
x_6	2	-6	0	$-\frac{5}{2}$	56	1	0	0
x_7	$\frac{1}{3}$	$-\frac{2}{3}$	0	$-\frac{1}{4}$	$\frac{16}{3}$	0	1	0
x_3	-2	6	1	$\frac{5}{2}$	-56	0	0	1
x_0	0	2	0	$\frac{7}{4}$	-44	$-\frac{1}{2}$	0	0
x_1	1	-3	0	$-\frac{5}{4}$	28	$\frac{1}{2}$	0	0
x_7	0	$\frac{1}{3}$	0	$\frac{1}{6}$	-4	$-\frac{1}{6}$	1	0
x_3	0	0	1	0	0	1	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	0	2	0	$\frac{7}{4}$	-44	$-\frac{1}{2}$	0	0
x_1	1	-3	0	$-\frac{5}{4}$	28	$\frac{1}{2}$	0	0
x_7	0	$\frac{1}{3}$	0	$\frac{1}{6}$	-4	$-\frac{1}{6}$	1	0
x_3	0	0	1	0	0	1	0	1

Example (cont)

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_0	0	2	0	$\frac{7}{4}$	-44	$-\frac{1}{2}$	0	0
x_1	1	-3	0	$-\frac{5}{4}$	28	$\frac{1}{2}$	0	0
x_7	0	$\frac{1}{3}$	0	$\frac{1}{6}$	-4	$-\frac{1}{6}$	1	0
x_3	0	0	1	0	0	1	0	1
x_0	0	0	0	$\frac{3}{4}$	-20	$\frac{1}{2}$	-6	0
x_1	1	0	0	$\frac{1}{4}$	-8	-1	9	0
x_2	0	1	0	$\frac{1}{2}$	-12	$-\frac{1}{2}$	3	0
x_3	0	0	1	0	0	1	0	1

This is the initial basic representation for $I = \{1, 2, 3\}$!

Bland's Rule

We can avoid cycling by amending the pivoting conventions.

Bland's Rule:

(i) Choose the lowest-numbered (leftmost) nonbasic column q with a positive cost.

$$q = \min\left\{j \neq 0 \mid \beta_j > 0\right\}$$

(ii) Choose the row $p \in \arg\min_{i \in I} \overline{x}_{iq}$ with the smallest index (same as standard conventions).

Theorem: With Bland's rule the simplex algorithm cannot cycle and hence is finite.

Degeneracy in Practice

- Until recently, cycling only occurred in contrived examples. It has been ignored in commercial codes.
- More recent experience with larger and larger problems indicates that cycling is considered a rare possibility.
- Rigorous remedies such as Bland's rule are not satisfactory as they
  - increase the number of iterations
  - and the work per iteration
  - in the majority of problems which would not cycle.
- In practice it is acceptable to replace a $y_{i0} = 0$ by $y_{i0} = \epsilon > 0$ (with $\epsilon = 10^{-2}$ or 10^{-3}) and then continue.