Infeasible IPMs

Infeasible IPMs

- SPF, PC, LPF algorithms start from a strictly feasible point (x^0, λ^0, s^0) in some neighborhood of the central path.
- Often, it is not easy to find a starting point that satisfies these conditions. Indeed, there are perfectly valid linear programs for which such points do not even exist.

Example. $\min\ 2x_1 + x_2$ s.t $x_1 + x_2 + x_3 = 5,\ x_1 + x_3 = 5,\ x \geq 0,$

For this problem, the primal feasible set is $\{(\beta, 0, 5-\beta) \mid \beta \in [0, 5]\}$.

Because $x_2 = 0$ for primal feasible set, the set $\mathcal{F}^o$ of strictly feasible points is empty.

Avoiding this difficulty

Avoiding this difficulty

- Homogeneous self-dual reformulation
- Infeasible-interior-point algorithm

Infeasible–interior point algorithm

- This algorithm does not require the initial point to be strictly feasible but requires only that it's x and s components be strictly.
- This algorithm works even when the set of strictly feasible points is empty

Infeasible IPMs

Definition: $$\mathcal{N}_{-\infty}(\gamma, \beta) = \{(x, \lambda, s) \mid \|(r_b, r_c)\| \le [\|(r_b^0, r_c^0)\|/\mu_0]\beta\mu, \quad (x, s) > 0, \quad x_i s_i \ge \gamma\mu, \; i = 1, 2, \dots, n\},$$

where $\gamma \in (0,1)$ and $\beta \ge 1$ are given parameters and (r_b^0, r_c^0) and μ_0 are evaluated at the starting point (x^0, λ^0, s^0).

- We note that we must have $\beta \ge 1$ to ensure that the initial point (x^0, λ^0, s^0) belongs to $\mathcal{N}_{-\infty}(\gamma, \beta)$.

Infeasible-interior-point algorithm

Given γ, β, $\sigma_{\min}$, $\sigma_{\max}$ with $\gamma \in (0,1)$, $\beta \geq 1$, and $0 < \sigma_{\min} < \sigma_{\max} \leq 0.5$;

choose (x^0, λ^0, s^0) with $(x^0, s^0) > 0$;

for $k = 0, 1, 2, \ldots$

choose $\sigma_k \in [\sigma_{\min}, \sigma_{\max}]$ and solve

$$\begin{bmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{bmatrix} \begin{bmatrix} \Delta x^k \\ \Delta \lambda^k \\ \Delta s^k \end{bmatrix} = \begin{bmatrix} -r_c^k \\ -r_b^k \\ -X^k S^k e + \sigma_k \mu_k e \end{bmatrix};$$

choose α_k as the largest value of α in $[0,1]$ such that

$$(x^k(\alpha), \lambda^k(\alpha), s^k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma, \beta)$$

and the Armijo condition $\mu_k(\alpha) \leq (1 - .01\alpha)\mu_k$; holds

set $(x^{k+1}, \lambda^{k+1}, s^{k+1}) = (x^k(\alpha_k), \lambda^k(\alpha_k), s^k(\alpha_k))$;

end(for)

Convergence of algorithm

$$r_b^k = Ax^k - b = A\left(x^{k-1} + \alpha_{k-1}\Delta x^{k-1}\right) - b$$
$$= r_b^{k-1} - \alpha_{k-1} r_b^{k-1} = (1 - \alpha_{k-1}) r_b^{k-1}$$
$$= (1 - \alpha_{k-1}) \dots (1 - \alpha_0) r_b^0$$

$$r_c^k = A^T(\lambda^{k-1} + \alpha_{k-1}\Delta\lambda^{k-1}) + s^{k-1} + \alpha_{k-1}\Delta s^{k-1} - c$$
$$= r_c^{k-1} - \alpha_{k-1} r_c^{k-1} = (1 - \alpha_{k-1}) r_c^{k-1}$$
$$= (1 - \alpha_{k-1}) \dots (1 - \alpha_0) r_c^0$$

- If $\alpha_k \leq 1$, then r_b^k and r_c^k are reduced in each iteration.

Convergence of algorithm

Theorem. The sequence $\{\mu_k\}$ and residual norms $\{\|(r_b^k, r_c^k)\|\}$ generated by infeasible algorithm converge to zero.

Theorem. Let $\epsilon > 0$ be given. Suppose that we use the starting point $(x^0, \lambda^0, s^0) = (\zeta e, 0, \zeta e)$,, where ζ is a scalar for which $\|(x^*, s^*)\|_\infty \leq \zeta$ for some primal-dual solution (x^*, λ^*, s^*). Suppose that the value of ζ for our particular linear programming instance satisfies $\zeta^2 \leq \frac{C}{\epsilon^\kappa}$ for some positive constants C and κ. Then there is an index K with

$$K = O\left(n^2 |\log \epsilon|\right)$$

such that the iterates $\{(x^k, \lambda^k, s^k)\}$ generated by Algorithm IPF satisfy

$$\mu_k \leq \epsilon \qquad \textit{for all } k \geq K.$$

Convergence of algorithm

Theorem. Let $\{(x^k, \lambda^k, s^k)\}$ be the sequence of iterates generated by Infeasible algorithm, and suppose that $\mathcal{F}^o \neq \emptyset$. Then partial sequence $\{(x^k, s^k)\}$ is bounded and therefore has at least one limit point

Theorem . Let $\{(x^k, \lambda^k, s^k)\}$ be the sequence of iterates generated by Infeasible algorithm. Then any limit point of the sequence $\{(x^k, s^k)\}$ can be used to construct a strictly complementary solution of primal and dual program.

Convergence of algorithm

- When the primal-dual feasible solution set $\mathcal{F}$, and hence the solution set Ω, is nonempty, an obvious stopping criterion for the algorithm is $\mu_k \leq \epsilon$, where ϵ is some small positive constant. Since the norm of the residual vector (r_c^k, r_b^k) is bounded by a multiple of μ_k , then the feasibility is also small and hence the current point is close to optimality. Therefore when a solution exists, infeasible algorithm will find it.

Theorem. If $\Omega = \emptyset$ the sequence $\{(x^k, \lambda^k, s^k)\}$ generated by infeasible algorithm has $\lim_{k\to\infty} \|(x^k, s^k)\| = \infty.$

Proof. We prove the result only for the case in which A has full row rank and leave the general case as an exercise. We reuse some of the analysis of Chapter 6 during this proof.

Suppose for contradiction that there is an infinite subsequence $\mathcal{K}$ and a constant ω such that

$$\|(x^k, s^k)\|_1 \le \omega \qquad \text{for all } k \in \mathcal{K}.$$

Because of compactness, there must be a subsequence $\bar{\mathcal{K}} \subset \mathcal{K}$ and a point $(\bar{x}, \bar{s})$ such that

$$\lim_{k \in \bar{\mathcal{K}}} (x^k, s^k) = (\bar{x}, \bar{s}).$$

Recall that μ_k is monotonically decreasing. If $\mu_k \downarrow 0$, we have from $(x^k, \lambda^k, s^k) \in \mathcal{N}_{-\infty}(\gamma, \beta)$ that $\|r_b^k\| \to 0$ and $\|r_c^k\| \to 0$. Hence for the limit point $(\bar{x}, \bar{s})$, we have $\bar{x}^T \bar{s} = 0$, $A\bar{x} = b$, and $A^T \bar{\lambda} + \bar{s} = c$ for some $\bar{\lambda}$. But then $(\bar{x}, \bar{\lambda}, \bar{s})$ is a primal-dual solution, which contradicts $\Omega = \emptyset$.

Hence, $\{\mu_k\}$ cannot approach zero, so there is a value $\bar{\mu} > 0$ such that $\mu_k \geq \bar{\mu}$ for all k. We define a set $\mathcal{G}^*$ that contains all such points by

$$\mathcal{G}^* = \left\{(x, \lambda, s) \in \mathcal{N}_{-\infty}(\gamma, \beta) \,|\, x^T s/n \geq \bar{\mu}, \;\; \|(x, s)\|_1 \leq \omega\right\}.$$

Note that $\mathcal{G}^*$ is compact and that $(x^k, \lambda^k, s^k) \in \mathcal{G}^*$ for all $k \in \mathcal{K}$. The solution $(\Delta x, \Delta\lambda, \Delta s)$ of the step equations (1.20) is unique in its Δx and Δs components for all $(x, \lambda, s) \in \mathcal{G}^*$ and all centering parameters $\sigma \in [0, 1]$. Therefore, by compactness, we can define a uniform upper bound $\hat{\omega} > 0$ on the step vector $(\Delta x, \Delta s)$ for all $(x, \lambda, s) \in \mathcal{G}^*$ and $\sigma \in [0, 1]$; that is,

$$\|(\Delta x^k, \Delta s^k)\| \leq \hat{\omega} \qquad \text{for all } k \in \mathcal{K} \text{ and some } \hat{\omega} > 0.$$

We can now use this step bound as the basis of a similar argument to the proof of Lemma 6.7, to show that there is a lower bound on the step length—a constant $\bar{\alpha} > 0$ such that $\alpha_k \geq \bar{\alpha}$ for all $k \in \mathcal{K}$. From the sufficient decrease condition (6.6), it follows that

$$\mu_{k+1} \leq (1 - .01\alpha_k)\mu_k \leq (1 - .01\bar{\alpha})\mu_k \qquad \text{for all } k \in \mathcal{K}.$$

Because $\{\mu_k\}$ is monotonic and $\mathcal{K}$ is an infinite sequence, it follows immediately that $\mu_k \downarrow 0$. This limit contradicts $\mu_k \geq \bar{\mu}$, so our proof is complete. □

Convergence of algorithm

- The previous theorem motivates the following simple termination test for algorithm IPF: for some positive constants ϵ (small) and ω (large),terminate the algorithm if

$$\mu_k \le \epsilon \qquad \text{or} \qquad \|(x^k, s^k)\|_1 \ge \omega. \qquad (2)$$

- If termination happens because of the first condition in (2), we are close to an approximation solution. If termination occurs because of the second condition in (2), chances are that Ω, and hence $\mathcal{F}$, is empty. Since ω is a finite number, we can not be certain that $\mathcal{F}$, is empty, but we can define a large set $\bar{\mathcal{F}}$ that does not intersect with $\mathcal{F}$.

Convergence of algorithm

- The set $\bar{\mathcal{F}}$ is defined in terms of two positive parameters $\bar{\delta}$ and $\bar{\omega}$ that are assumed to satisfy the condition

$$\frac{\bar{\omega}^2 + n\mu_0}{\bar{\delta}} \leq \omega. \qquad (9.11)$$

$$\bar{\mathcal{F}} = \{(\bar{x}, \bar{\lambda}, \bar{s}) \mid \bar{x} \geq \bar{\delta}e, \ \ \bar{s} \geq \bar{\delta}e, \ \ \|(\bar{x}, \bar{s})\|_1 \leq \bar{\omega}\}$$

Theorem. Suppose that the starting point for Algorithm IPF lies Inside $\bar{\mathcal{F}}$ and that the algorithm terminates at some iterate $k > 0$ with $\|(x^k, s^k)\|_1 \geq \omega$. Then $\bar{\mathcal{F}}$ contains no primal-dual feasible vectors, that is, $\bar{\mathcal{F}} \cap \mathcal{F} = \emptyset$.

Proof. We prove the result by contradiction, showing that if there exists a feasible point inside $\bar{\mathcal{F}}$, the condition (9.11) must be violated.

Recall that for ν_k defined by (6.7) and r_b and r_c defined by (6.1), we have from (6.8) that

$$Ax^k - b = r_b^k = \nu_k r_b^0, \qquad A^T\lambda^k + s^k - c = r_c^k = \nu_k r_c^0. \tag{9.13}$$

Suppose that we have a point (x, λ, s) that lies in $\mathcal{F} \cap \bar{\mathcal{F}}$, and define another point $(\bar{x}, \bar{\lambda}, \bar{s})$ by

$$(\bar{x}, \bar{\lambda}, \bar{s}) = (1 - \nu_k)(x, \lambda, s) + \nu_k(x^0, \lambda^0, s^0).$$

Since (x, λ, s) and (x^0, λ^0, s^0) both belong to the convex set $\bar{\mathcal{F}}$, so does $(\bar{x}, \bar{\lambda}, \bar{s})$, since it lies on a line between these two points.

Because of (9.13) and feasibility of (x, λ, s), we have that

$$A\bar{x} - b = \nu_k r_b^0, \qquad A^T\bar{\lambda} + \bar{s} - c = \nu_k r_c^0$$

so that $(\bar{x}, \bar{\lambda}, \bar{s})$ satisfies the same equations as (x^k, λ^k, s^k) in (9.13). Hence, by a standard argument (see, for instance, the proof of Lemma 2.5), we have

$$(\bar{x} - x^k)^T(\bar{s} - s^k) = 0. \tag{9.14}$$

Therefore we have

$$
\begin{aligned}
\bar{\omega}^2 + n\mu_0 &\geq \bar{x}^T \bar{s} + (x^k)^T s^k && \text{since } \|(\bar{x}, \bar{s})\|_1 \leq \bar{\omega} \text{ and } (x^k)^T s^k < n\mu_0 \\
&= \bar{x}^T s^k + \bar{s}^T x^k && \text{by (9.14)} \\
&\geq \bar{\delta} e^T s^k + \bar{\delta} e^T x^k && \text{since } \bar{x} \geq \delta e \text{ and } \bar{s} \geq \delta e \\
&= \bar{\delta} \|(x^k, s^k)\|_1 && \\
&\geq \bar{\delta} \omega && \text{since } \|(x^k, s^k)\|_1 \geq \omega \text{ at termination.}
\end{aligned}
$$

Since this inequality contradicts (9.11), we conclude that $\mathcal{F} \cap \bar{\mathcal{F}}$ must be empty. □

Starting point

- Global convergence of infeasible algorithm is guaranteed from any starting point (x^0, λ^0, s^0) with $(x^0, s^0) > 0$. However a good starting point should also satisfy two other conditions. First the points should be well centered, so that pairwise products $x_i^0 s_i^0$ are similar for all $i = 1, 2, \ldots, n$. Second, the point should not be too infeasible, that is, the ratio $\|(r_b^0, r_c^0)\|/\mu_0$ of infeasibility to duality measure should not be too large.
- A popular heuristic for finding (x^0, λ^0, s^0) starts by calculating $(\tilde{x}, \tilde{\lambda}, \tilde{s})$ as the solution of two least-squares problems:

$$\min_x \|x\|^2 \quad \text{subject to } Ax = b,$$
$$\min_{(\lambda,s)} \|s\|^2 \quad \text{subject to } A^T\lambda + s = c.$$

That is, $\tilde{x}$ and $\tilde{s}$ are the vectors of least norm for which the residuals r_b

Starting point

and r_c are zero. The starting point then is defined as:

$$(x^0, \lambda^0, s^0) = (\tilde{x} + \tilde{\delta}_x e, \tilde{\lambda}, \tilde{s} + \tilde{\delta}_s e),$$

where

$$\bar{\delta_x} = \delta_x + 0.5 * \frac{(\bar{x} + \delta_x e)^T (\bar{s} + \delta_x e)}{\sum_{i=1}^{n} (s_i + \delta_x)}$$

$$\bar{\delta_s} = \delta_s + 0.5 * \frac{(\bar{x} + \delta_x e)^T (\bar{s} + \delta_x e)}{\sum_{i=1}^{n} (x_i + \delta_x)}$$

$$\delta_x = \max\left(-1.5 * \min\{\bar{x}_i\}, 0\right)$$

$$\delta_s = \max\left(-1.5 * \min\{\bar{s}_i\}, 0\right)$$