COMP 545: Advanced topics in optimization
From simple to complex ML systems

Lecture 4

Line of thought towards IPMs

- Lagrangian methods describe conditions that lead to optimal solutions of constrained problems.

Line of thought towards IPMs

- Lagrangian methods describe conditions that lead to optimal solutions of constrained problems.

- These conditions, when constraints are simple enough, lead to closed−form solutions (by solving e.g., a system of linear equations).

Line of thought towards IPMs

- Lagrangian methods describe conditions that lead to optimal solutions of constrained problems.

- These conditions, when constraints are simple enough, lead to closed–form solutions (by solving e.g., a system of linear equations).

- Further, if constraints are simple enough, one can form augmented Lagrangian and solve the problem with faster convergence.

Line of thought towards IPMs

- Lagrangian methods describe conditions that lead to optimal solutions of constrained problems.

- These conditions, when constraints are simple enough, lead to closed–form solutions (by solving e.g., a system of linear equations).

- Further, if constraints are simple enough, one can form augmented Lagrangian and solve the problem with faster convergence.

- There are problems though that lead conditions that do not have closed–form solutions, neither there is an easy augmented Lagrangian formulation.

Line of thought towards IPMs

- Lagrangian methods describe conditions that lead to optimal solutions of constrained problems.
- These conditions, when constraints are simple enough, lead to closed–form solutions (by solving e.g., a system of linear equations).
- Further, if constraints are simple enough, one can form augmented Lagrangian and solve the problem with faster convergence.
- There are problems though that lead conditions that do not have closed–form solutions, neither there is an easy augmented Lagrangian formulation.
- One can model problems with barriers into the objective to model the constraints.

Line of thought towards IPMs

- Constrained problems though often have optimal solutions on the boundary of the constraint set.

$$-\log(a_i^\top x - b_i)$$

Line of thought towards IPMs

- Constrained problems though often have optimal solutions on the boundary of the constraint set.

$$-\log(a_i^\top x - b_i)$$

- So, solving these problems using powerful routines, like Newton's method, gets us into "trouble", as the barrier terms explode (this is reflected into the condition number of the problem).

Line of thought towards IPMs

- Constrained problems though often have optimal solutions on the boundary of the constraint set.

$$-\log(a_i^\top x - b_i)$$

- So, solving these problems using powerful routines, like Newton's method, gets us into "trouble", as the barrier terms explode (this is reflected into the condition number of the problem).

- Newton's method (at least the analysis) heavily depends on the condition number. We need analysis that takes into account this singularity.

Line of thought towards IPMs

- Constrained problems though often have optimal solutions on the boundary of the constraint set.

$$-\log(a_i^\top x - b_i)$$

- So, solving these problems using powerful routines, like Newton's method, gets us into "trouble", as the barrier terms explode (this is reflected into the condition number of the problem).

- Newton's method (at least the analysis) heavily depends on the condition number. We need analysis that takes into account this singularity.

- We need to define a different distance norm that takes into account this.

$$\|h\|_x \equiv \sqrt{h^\top \nabla^2 f(x) h}$$

Line of thought towards IPMs

- This norm depends on the Hessian; so if the Hessian changes quickly, we might not be able to bound using

$$\|x - x^\star\|_x$$

Line of thought towards IPMs

- This norm depends on the Hessian; so if the Hessian changes quickly, we might not be able to bound using

$$\|x - x^\star\|_x$$

- Thus, the rate of change of the Hessian should be bounded.

Line of thought towards IPMs

- This norm depends on the Hessian; so if the Hessian changes quickly, we might not be able to bound using

$$\|x - x^\star\|_x$$

- Thus, the rate of change of the Hessian should be bounded.
- This reasoning leads us to impose a bound on the third derivative in terms of the Hessian:

$$|f'''(x)| \leq 2f''(x)^{3/2}$$

Line of thought towards IPMs

- This norm depends on the Hessian; so if the Hessian changes quickly, we might not be able to bound using

$$\|x - x^\star\|_x$$

- Thus, the rate of change of the Hessian should be bounded.
- This reasoning leads us to impose a bound on the third derivative in terms of the Hessian:

$$|f'''(x)| \leq 2f''(x)^{3/2}$$

- We need to define a different distance norm that takes into account this.

$$\|h\|_x \equiv \sqrt{h^\top \nabla^2 f(x) h} \qquad |\nabla^3 f(x)[h, h, h]| \leq 2\|h\|_x^3$$

Figure 1

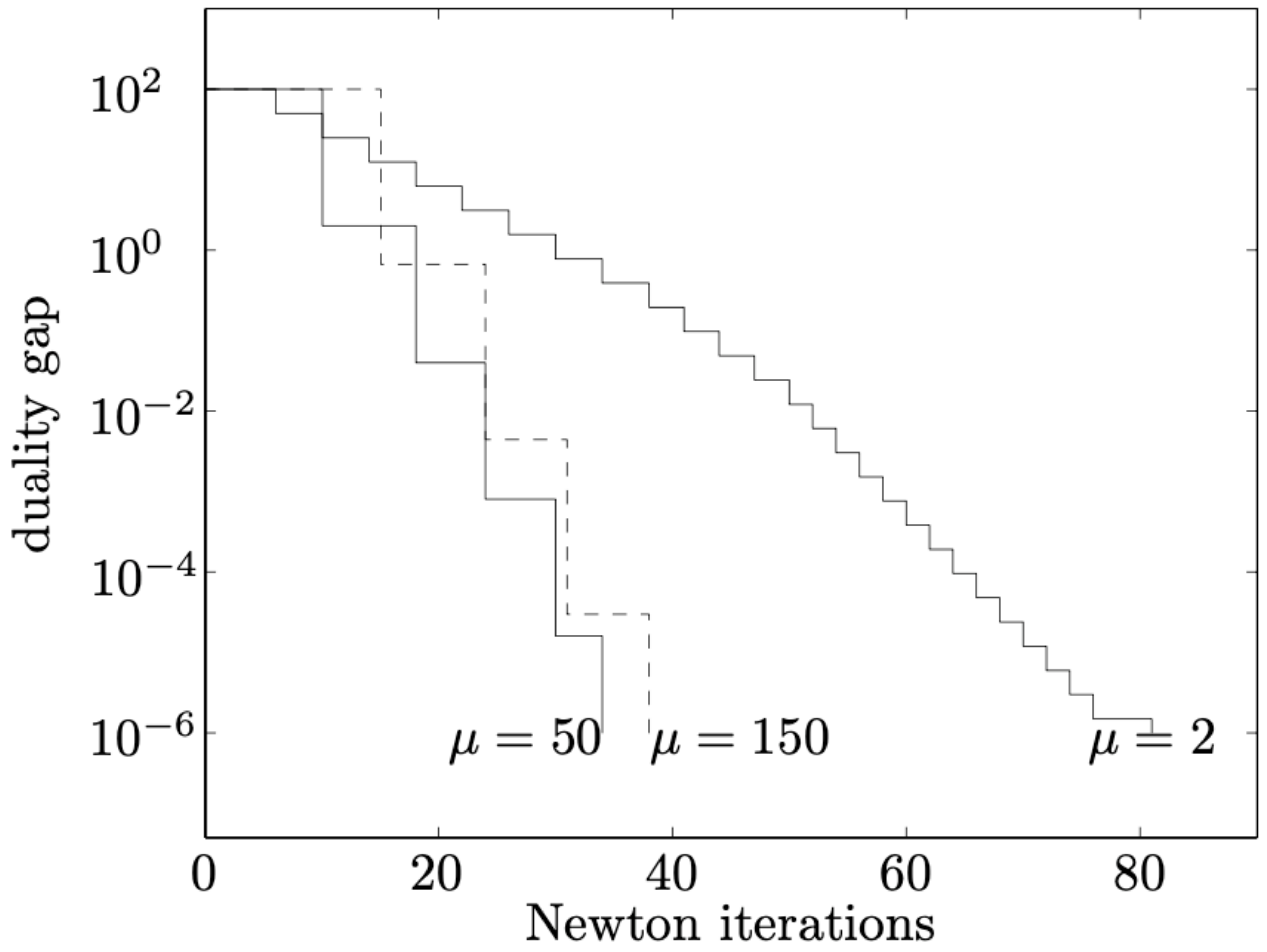

Figure 11.4 Progress of barrier method for a small LP, showing duality gap versus cumulative number of Newton steps. Three plots are shown, corresponding to three values of the parameter μ: 2, 50, and 150. In each case, we have approximately linear convergence of duality gap.

Figure 2

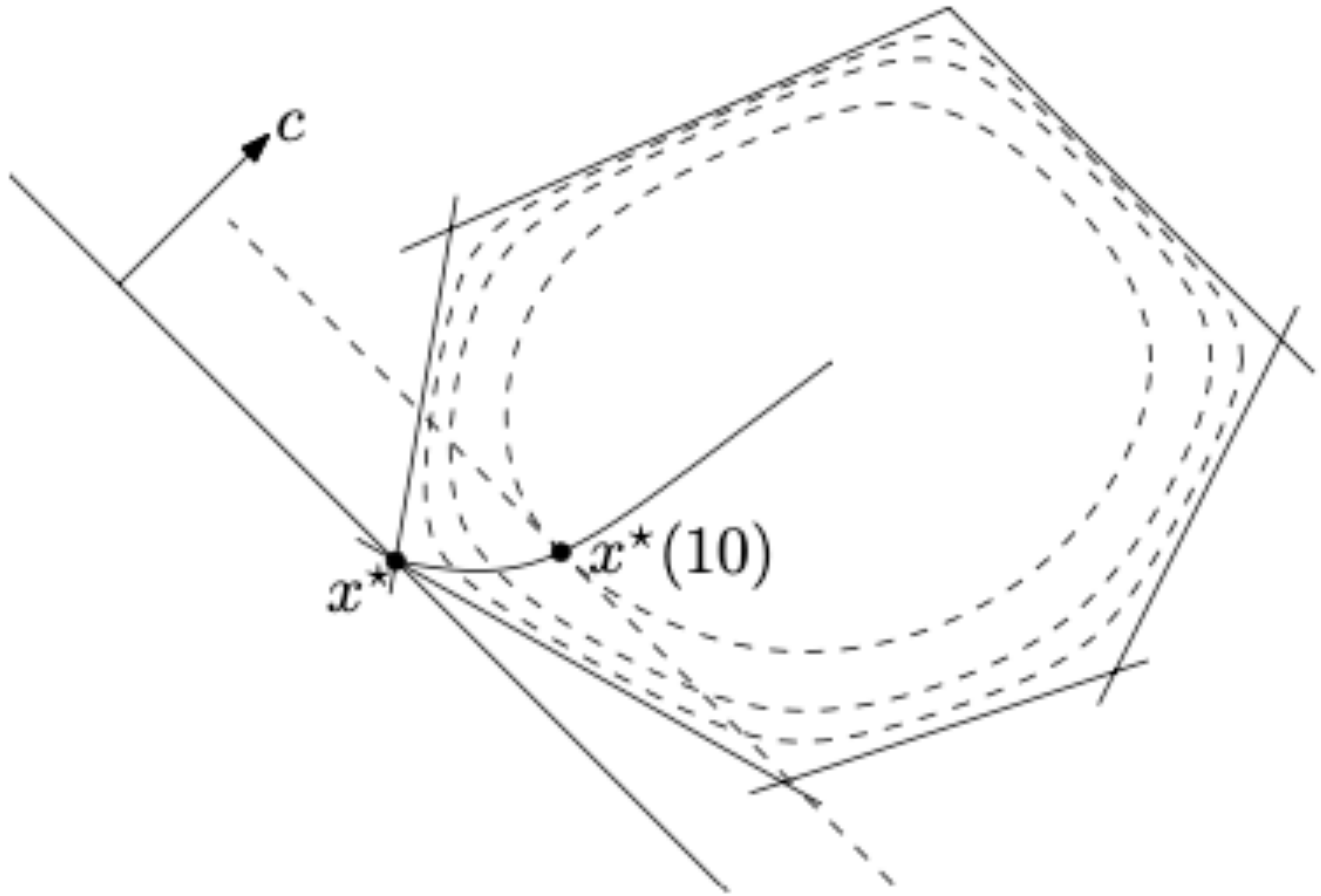

Figure 11.2 Central path for an LP with $n = 2$ and $m = 6$. The dashed curves show three contour lines of the logarithmic barrier function ϕ. The central path converges to the optimal point $x^\star$ as $t \to \infty$. Also shown is the point on the central path with $t = 10$. The optimality condition (11.9) at this point can be verified geometrically: The line $c^T x = c^T x^\star(10)$ is tangent to the contour line of ϕ through $x^\star(10)$.