Warmstarting Interior-Point Methods

.

Hande Y. Benson

Drexel University

David F. Shanno, Rutgers University - RUTCOR

Warmstarting

- Problem data/characteristics may be subject to change
  - Market prices, demand
  - New products, new restrictions
  - Environmental conditions, design specifications
  - Integer programming algorithms
- It may be necessary to solve a series of closely related optimization problems.
- *Warmstarting* is the use of information obtained during the solution of the initial problem to solve the subsequent perturbed problems.

Linear Programming

Warmstarting LPs

Consider the Linear Programming Problem (LP) of the form:

$$\begin{array}{lll} \max_{x,w} & c^T x & \\ \text{s.t.} & Ax + w & = & b \\ & x, w & \geq & 0 \end{array}$$

Its dual has the form:

$$\begin{array}{lll} \min_{y,z} & b^T y & \\ \text{s.t.} & A^T y - z & = & c \\ & y, z & \geq & 0 \end{array}$$

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Interior-Point Methods for LP

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Optimality conditions:

$$\begin{aligned} Ax + w &= b \\ A^T y - z &= c \\ x_j z_j &= 0, \quad j = 1, \dots, n \\ w_i y_i &= 0, \quad i = 1, \dots, m, \end{aligned}$$

Relax the complementarity conditions:

$$\begin{aligned} x_j z_j &= \mu, \quad j = 1, \dots, n \\ w_i y_i &= \mu, \quad i = 1, \dots, m, \end{aligned}$$

Interior-Point Methods for LP - II

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Apply Newton's Method, solve the reduced KKT system at each iteration:

$$\begin{bmatrix} -E & A \\ A^T & F \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \end{bmatrix} = \begin{bmatrix} \rho - E\gamma_w \\ \sigma + \gamma_z \end{bmatrix},$$

where

$$\begin{aligned} E &= Y^{-1}W & F &= X^{-1}Z \\ \rho &= b - Ax - w & \sigma &= c - A^T y + z \\ \gamma_w &= \mu W^{-1}e - y & \gamma_z &= \mu X^{-1}e - z. \end{aligned}$$

We also have that

$$\begin{aligned} \Delta z &= \gamma_z - F\Delta x \\ \Delta w &= E(\gamma_w - \Delta y). \end{aligned}$$

An Example

$$\begin{array}{lrcl} \text{maximize} & x_1 + 3x_2 & & \\ \text{subject to} & x_1 + x_2 + w_1 & = & 3 \\ & 2x_1 + x_2 + w_2 & = & 2 \\ & x_1, x_2, w_1, w_2 & \geq & 0. \end{array}$$

The optimal primal and dual solutions for this problem are

$$\begin{array}{rclcrcl} x & = & (0,2), & & w & = & (1,0) \\ y & = & (0,3), & & z & = & (5,0). \end{array}$$

At the optimum, the first constraint is inactive and the second constraint is active.

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Perturbing the Example

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

$$\begin{array}{lrcl} \text{maximize} & x_1 + 3x_2 & & \\ \text{subject to} & x_1 + x_2 + w_1 & = & 1 \\ & 2x_1 + x_2 + w_2 & = & 2 \\ & x_1, x_2, w_1, w_2 & \geq & 0. \end{array}$$

The optimal primal and dual solutions for this problem are

$$\begin{array}{rclrcl} x & = & (0,1), & w & = & (0,1) \\ y & = & (3,0), & z & = & (2,0). \end{array}$$

At the optimum, the first constraint is active and the second constraint is inactive .

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Not warm enough....

Problems:

- The reduced KKT system:

$$-\frac{w_1}{y_1}\Delta y_1 + \Delta x_1 + \Delta x_2 = \rho_1 - \frac{w_1}{y_1}\gamma_{w1}.$$

- The barrier parameter:

$$\mu = r\frac{w^T y + x^T z}{m+n}$$

- Iterative refinement

Iteration	x	w	y	z
0	(0,2)	(1,0)	(0,3)	(5,0)
1	(0,2)	(0.05,0)	(0,3)	(5,0)
2	(0,2)	(0.0025,0)	(0,3)	(5,0)
3*	(0,2)	(0.000125,0)	(1.6,1.4)	(3.4,0)
4*	(0,2)	(0.00000625,0)	(2.93,0.07)	(2.07,0)

It takes 15 iterations for LOQO to find the optimal solution.
The coldstart takes 12!

Primal-Dual Penalty Model

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

The primal problem

$$
\begin{array}{rrlcl}
\max\limits_{x,w,\xi_x,\xi_w} & & c^T x - d_x^T \xi_x - d_w^T \xi_w & & \\
\text{s.t.} & & Ax + w & = & b \\
-\xi_x & \leq & x & \leq & u_z \\
-\xi_w & \leq & w & \leq & u_y \\
& & \xi_x, \xi_w & \geq & 0,
\end{array}
$$

and the corresponding dual

$$
\begin{array}{rrlcl}
\min\limits_{y,z,\psi_y,\psi_z} & & b^T y + u_y^T \psi_y + u_z^T \psi_z & & \\
\text{s.t.} & & A^T y - z & = & c \\
-\psi_y & \leq & y & \leq & d_w - \psi_y \\
-\psi_z & \leq & z & \leq & d_x - \psi_z \\
& & \psi_y, \psi_z & \geq & 0.
\end{array}
$$

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Numerical Testing Details

- 37 problems from the Netlib test suite
- Solved using LOQO Version 6.06
- Uniformly distributed random numbers between -1 and 1
- Update no more than 10% or 20 data elements (b, c, or A)

$$\tilde{b}_i = \begin{cases} \delta\epsilon, & \text{if } b_i = 0 \\ b_i(1+\delta\epsilon), & \text{otherwise.} \end{cases}$$

- Repeated for $\delta = 0.001, 0.01,$ and 0.1.

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Perturbing b

Problem	δ	#Pert	Warm	Cold	$\|x^I - x^P\|$	$\|y^I - y^P\|$	Δ
adlittle	0.001	5.6	2.6	18	3.20e-3	1.46e-5	2
adlittle	0.01	5.6	3.8	17.8	3.64e-2	1.15e-2	4
adlittle	0.1	5.6	11	17	3.92e-1	1.81e0	7.8
agg3	0.001	20.2	5	24	4.44e-4	5.21e-4	0.2
agg3	0.01	20.2	6.6	24	4.97e-3	2.52e-1	11.6
agg3	0.1	20.2	9.6	24	4.46e-2	9.25e-1	23.4
bandm	0.001	20.2	10.4	19	1.24e-4	1.09e0	6.2
bandm	0.01	20.2	10.6	19.2	1.24e-3	9.26e-1	0.8
bandm	0.1	20.2	11.2	20	1.10e-2	9.24e-1	3.4

Changed	δ	Avg Warmstart Iters	Avg Coldstart Iters	Avg Reduction
b	0.001	10.38	21.76	52%
b	0.01	10.41	21.67	52%
b	0.1	12.07	21.58	44%

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Perturbing c

Problem	δ	#Pert	Warm	Cold	$\|x^I - x^P\|$	$\|y^I - y^P\|$	Δ
adlittle	0.001	9	5	18.2	1.36e-1	1.64e-2	7.4
adlittle	0.01	9	6.2	18.2	1.45e-1	3.95e-2	0
adlittle	0.1	9	8.2	16.4	2.24e-1	1.21e-1	1
agg2	0.001	20.6	26.2	53.6	9.47e-2	3.02e-2	15.6
agg2	0.01	20.6	12.2	26.6	1.83e-1	2.83e0	2
agg2	0.1	20.6	13.8	26.6	3.13e-1	5.08e0	28.8
bandm	0.001	20.2	10	19	2.25e-7	9.73e-1	1.8
bandm	0.01	20.2	9.8	19	4.12e-3	1.13e0	3.2
bandm	0.1	20.2	11.8	19.6	1.41e-2	6.50e-1	5.8

Changed	δ	Avg Warmstart Iters	Avg Coldstart Iters	Avg Reduction
c	0.001	11.55	22.74	49%
c	0.01	11.78	22.09	47%
c	0.1	13.00	22.21	41%

Perturbing A

Introduction

Linear Programming
- Warmstarting LPs
- Interior-Point Methods for LP
- An Example
- Primal-Dual Penalty Model
- Numerical Testing Details
- Perturbing b
- Perturbing c
- Perturbing A

Nonlinear Programming

More Details

Problem	δ	#Pert	Warm	Cold	$\|\|x^I - x^P\|\|$	$\|\|y^I - y^P\|\|$	Δ
adlittle	0.001	20	5.8	18.8	2.30e-1	3.07e-2	4
adlittle	0.01	20	7.4	18.8	2.42e-1	7.94e-2	12.2
adlittle	0.1	20	10.4	17.4	2.61e-1	2.37e-1	2.2
bandm	0.001	18	10	19	1.68e-4	9.70e-1	2.8
bandm	0.01	18	10.6	19.2	2.26e-3	1.01e0	0.2
bandm	0.1	18	12	19	2.39e-2	1.37e0	5.2
beaconfd	0.001	18.2	3	14	1.08e-3	7.78e-4	8.2
beaconfd	0.01	18.2	3.6	14	3.04e-2	2.88e-2	8.4
beaconfd	0.1	18.2	6.6	14.2	2.03e-1	1.83e0	4.2

Changed	δ	Avg Warmstart Iters	Avg Coldstart Iters	Avg Reduction
A	0.001	12.01	21.78	45%
A	0.01	12.41	21.78	43%
A	0.1	13.95	21.87	36%

Nonlinear Programming

Introduction

Linear Programming

Nonlinear Programming
- Warmstarting NLPs
- Interior-Point Methods for NLP
- Primal-Dual Penalty Model - NLP
- Nonlinear Programming: Real World Problems
- Quadratic Programming
- Hock and Schittkowski Test Suite

More Details

Warmstarting NLPs

Consider the Nonlinear Programming Problem (NLP) of the form:

$$\begin{array}{ll} \min\limits_{x,w} & f(x) \\ \text{s.t.} & h(x) - w = 0 \\ & w \geq 0, \end{array}$$

Its dual has the form:

$$\begin{array}{ll} \max\limits_{y} & f(x) - \nabla f(x)^T x - (h(x) - A(x)x)^T y \\ \text{s.t.} & \nabla f(x) - A(x)^T y = 0 \\ & y \geq 0, \end{array}$$

Interior-Point Methods for NLP

Introduction

Linear Programming

Nonlinear Programming

- Warmstarting NLPs
- Interior-Point Methods for NLP
- Primal-Dual Penalty Model - NLP
- Nonlinear Programming: Real World Problems
- Quadratic Programming
- Hock and Schittkowski Test Suite

More Details

First-order conditions:

$$\nabla f(x) - A(x)^T y = 0$$

$$h(x) - w = 0$$

$$WYe = 0,$$

Relax the complementarity conditions:

$$WYe \quad = \quad \mu e.$$

The reduced KKT system is solved at each iteration:

$$\begin{bmatrix} -H(x,y) & A(x)^T \\ A(x) & WY^{-1} \end{bmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} \nabla f(x) - A^T(x)y \\ -h(x) + \mu Y^{-1} e \end{pmatrix},$$

Primal-Dual Penalty Model - NLP

Introduction

Linear Programming

Nonlinear Programming

- Warmstarting NLPs
- Interior-Point Methods for NLP
- Primal-Dual Penalty Model - NLP
- Nonlinear Programming: Real World Problems
- Quadratic Programming
- Hock and Schittkowski Test Suite

More Details

The primal problem

$$
\begin{aligned}
\min_{x,w,\xi} \quad & f(x) + d^T \xi \\
\text{s.t.} \quad & h(x) - w = 0 \\
& -\xi \leq w \leq u \\
& \xi \geq 0,
\end{aligned}
$$

and the corresponding dual

$$
\begin{aligned}
\max_{y,\psi} \quad & f(x) - \nabla f(x)^T x - (h(x) - A(x)x)^T y - u^T \psi \\
\text{s.t.} \quad & \nabla f(x) - A(x)^T y = 0 \\
& -\psi \leq y \leq d - \psi \\
& \psi \geq 0,
\end{aligned}
$$

Nonlinear Programming: Real World Problems

- Randomly perturbed problem parameters
- "*" indicates nonconvex problem

Problem	Avg Warmstart Iters	Avg Coldstart Iters	Avg Reduction
antenna*	34.0	27.8	-24.28%
blend*	11.2	20.6	45.44%
chemeq	40.2	299.2	86.46%
fi r_convex	37.8	42.0	8.61%
hang_midpt*	9.6	45.8	78.96%
hydrothermal*	28.2	28.0	-0.71%
kowalik*	7.4	11.0	32.73%
markowitz	7.4	14.0	47.14%
nb_L2	6.2	23.0	72.82%
nnls	7.6	13.0	41.54%
robotarm*	19.6	22.8	14.07%
rocket*	32.8	57.4	36.72%
sawpath*	2.8	11.0	74.55%
shekel*	5.2	18.6	70.30%
springs	5.0	59.4	89.76%
steiner	17.8	23.0	23.75%
wbv	11.8	17.6	28.86%

Introduction

Linear Programming

Nonlinear Programming
- Warmstarting NLPs
- Interior-Point Methods for NLP
- Primal-Dual Penalty Model - NLP
- Nonlinear Programming: Real World Problems
- Quadratic Programming
- Hock and Schittkowski Test Suite

More Details

Quadratic Programming

Introduction

Linear Programming

Nonlinear Programming
- Warmstarting NLPs
- Interior-Point Methods for NLP
- Primal-Dual Penalty Model - NLP
- Nonlinear Programming: Real World Problems
- Quadratic Programming
- Hock and Schittkowski Test Suite

More Details

- Convex QPs from the CUTEr test set
- Perturbed b, c, A as in LP

Changed	δ	Avg Warmstart Iters	Avg Coldstart Iters	Avg Reduction
b	0.0001	24.63	53.54	45%
b	0.001	34.40	68.48	49%
b	0.01	35.97	66.63	42%
c	0.0001	10.89	39.80	49%
c	0.001	10.29	39.07	49%
c	0.01	10.08	39.51	48%
A	0.0001	17.19	53.91	55%
A	0.001	16.89	54.11	55%
A	0.01	18.11	54.64	49%

Introduction

Linear Programming

Nonlinear Programming
- Warmstarting NLPs
- Interior-Point Methods for NLP
- Primal-Dual Penalty Model - NLP
- Nonlinear Programming: Real World Problems
- Quadratic Programming
- Hock and Schittkowski Test Suite

More Details

Hock and Schittkowski Test Suite

Perturb the optimal solution, then resolve.

δ	Reduction in Iters	Primal Dist	Dual Dist
0.0001	39%	3.70e-03	2.13e-03
0.001	33%	6.14e-03	1.72e-02
0.01	24%	1.90e-02	3.40e-02

- Average performance of the primal-dual penalty approach when warmstarting 103 problems from the Hock and Schittkowski test set
- δ is the perturbation factor
- The second column is the average reduction in iterations from the coldstart to the warmstart
- Primal Dist and Dual Dist are the scaled Euclidean distances between the warmstart and the optimal solutions for the primal and the dual variables, respectively.

More Details

Exactness of the Penalty Model

Set the penalty parameters so that

$$u > \bar{w} \text{ and } d > \bar{y}.$$

Then, the optimality conditions

$$\begin{aligned} \nabla f(x) - A(x)^T y &= 0 \\ h(x) - w &= 0 \\ \mu e - (W + \Xi)(Y + \Psi)e &= 0 \\ \mu e - \Xi(D - Y - \Psi)e &= 0 \\ \mu e - \Psi(U - W)e &= 0 \end{aligned}$$

reduce to the optimality conditions of the original problem.

Introduction

Linear Programming

Nonlinear Programming

More Details
- Exactness of the Penalty Model
- The Reduced KKT System
- Problems in General Form
- Computational Issues: Initialization
- Computational Issues: Updates
- Penalty Methods for General NLP
- General Nonlinear Programming

The Reduced KKT System

$$\begin{bmatrix} -H(x,y) & A(x)^T \\ A(x) & E \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} \sigma \\ \rho + E\gamma \end{bmatrix}$$

where

$$\begin{aligned} E &= \left(\left((Y+\Psi)^{-1}(W+\Xi) + \Xi(D-Y-\Psi)^{-1} \right)^{-1} + \Psi(U-W)^{-1} \right)^{-1} \\ \gamma &= \left((Y+\Psi)^{-1}(W+\Xi) + \Xi(D-Y-\Psi)^{-1} \right)^{-1} \\ & \quad \left(\mu(Y+\Psi)^{-1}e - \mu(D-Y-\Psi)^{-1}e - w \right) - \left(\mu(U-W)^{-1}e - \psi \right) \end{aligned}$$

with

$$\begin{aligned} \Delta w &= -(W+\Xi)(Y+\Psi)^{-1}(\gamma_w + \Delta y + \Delta\psi) - \Delta\xi, \\ \Delta\xi &= (D-Y-\Psi)^{-1}\Xi(\Delta y + \Delta\psi) - \gamma_\xi, \\ \Delta\psi &= -\left((Y+\Psi)^{-1}(W+\Xi) + (U-W)\Psi^{-1} + (D-Y-\Psi)^{-1}\Xi \right)^{-1} \\ & \quad \left(u - \mu\Psi^{-1}e - \mu(Y+\Psi)^{-1}e + \mu(D-Y-\Psi)^{-1}e \right) \\ & \quad - \left((Y+\Psi)^{-1}(W+\Xi) + (U-W)\Psi^{-1} + (D-Y-\Psi)^{-1}\Xi \right)^{-1} \\ & \quad \left((Y+\Psi)^{-1}(W+\Xi) + (D-Y-\Psi)^{-1}\Xi \right) \Delta y \end{aligned}$$

Introduction

Linear Programming

Nonlinear Programming

More Details
- Exactness of the Penalty Model
- The Reduced KKT System
- Problems in General Form
- Computational Issues: Initialization
- Computational Issues: Updates
- Penalty Methods for General NLP
- General Nonlinear Programming

Problems in General Form

Introduction
Linear Programming
Nonlinear Programming
More Details
- Exactness of the Penalty Model
- The Reduced KKT System
- Problems in General Form
- Computational Issues: Initialization
- Computational Issues: Updates
- Penalty Methods for General NLP
- General Nonlinear Programming

- A variable with the bounds

$$x_l \le x \le x_u,$$

is converted to

$$\begin{aligned} x - g &= x_l \\ x + t &= x_u \\ g, t &\ge 0. \end{aligned}$$

- A range constraint of the form

$$0 \le h(x) \le r$$

is converted to

$$\begin{aligned} h(x) - w &= 0 \\ w + p &= r \\ w, p &\ge 0. \end{aligned}$$

Computational Issues: Initialization

Introduction

Linear Programming

Nonlinear Programming

More Details
- Exactness of the Penalty Model
- The Reduced KKT System
- Problems in General Form
- Computational Issues: Initialization
- Computational Issues: Updates
- Penalty Methods for General NLP
- General Nonlinear Programming

How do we initialize the relaxation variables and the penalty parameters in order to reach the new optimum quickly after a warmstart?

Relaxation variables:

$$\begin{aligned} \xi &= \max(h(x) - w, 0) + \tau \\ \psi &= \tau, \end{aligned}$$

where τ is a small parameter, currently set to $10^{-5}M$, where M is the greater of 1 and the largest primal or dual slack value.

Penalty parameters:

$$\begin{aligned} u &= 10(w + \kappa) \\ d &= 10(y + \psi + \kappa), \end{aligned}$$

where κ is a constant with a default value of 1.

Introduction
Linear Programming
Nonlinear Programming
More Details
- Exactness of the Penalty Model
- The Reduced KKT System
- Problems in General Form
- Computational Issues: Initialization
- Computational Issues: Updates
- Penalty Methods for General NLP
- General Nonlinear Programming

Computational Issues: Updates

If necessary, how do we update the penalty parameters in order to obtain the solution to the perturbed LP?

- Static updates
- Dynamic updates

If $w_i^{(k)} > 0.9u_i^{(k)}$, then $u_i^{(k+1)} = 10u_i^{(k)}$, $i = 1, \ldots, m.$

If $y_i^{(k)} + \psi_i^{(k)} > 0.9d_i^{(k)}$, then $d_i^{(k+1)} = 10d_i^{(k)}$, $i = 1, \ldots, m.$

Introduction

Linear Programming

Nonlinear Programming

More Details
- Exactness of the Penalty Model
- The Reduced KKT System
- Problems in General Form
- Computational Issues: Initialization
- Computational Issues: Updates
- Penalty Methods for General NLP
- General Nonlinear Programming

Penalty Methods for General NLP

- Benefits include:
  - Primal and dual infeasibility detection
  - Bounded sets of optimal primal and dual solutions
  - Relieving of the jamming phenomenon
  - Detection of nonKKT optima

Introduction

Linear Programming

Nonlinear Programming

More Details
- Exactness of the Penalty Model
- The Reduced KKT System
- Problems in General Form
- Computational Issues: Initialization
- Computational Issues: Updates
- Penalty Methods for General NLP
- General Nonlinear Programming

General Nonlinear Programming

- 1078 models from the CUTEr set, written in AMPL
- ℓ_1 penalty method implemented in LOQO Version 6.06 (current distribution)

	Problems Solved	Total Iterations	Total Runtime
Default	986	50520	3029.73
Penalty	977	60820	7096.78

	Problems Solved	Total Iterations	Total Runtime
Default	986	57116	3927.50
Hybrid	1002	39777	4881.23

	Problems Solved	Total Iterations
Hybrid LOQO	1016	55819
IPOPT	1013	39061