

# Lecture XIX: Linearly Constrained Algorithm and Nonlinear Algorithms

Charles B. Moss<sup>1</sup>

<sup>1</sup>University of Florida

October 4, 2012

- 1 Linear Equality Constraints
  - Maximize Utility Subject to a Budget Constraint
  - Starting Point
  - Compute Second Step
  - Compute the Third Step
  - Compute the Fourth Step
  
- 2 Algorithms for Nonlinear Constraints
  - Penalty Functions

- As a starting point, we assume a standard Cobb-Douglas Utility function specification subject to an income constraint

$$\begin{aligned} \max_{x_1, x_2, x_3} \quad & U(x_1, x_2, x_3) = x_1^{0.2} x_2^{0.3} x_3^{0.4} \\ \text{s.t.} \quad & Ax \leq b \Rightarrow x_1 + 2x_2 + 3x_3 \leq 100 \end{aligned} \quad (1)$$

- Compute the direction of search

$$p_1 = -\nabla_x U(x) [\nabla_{xx}^2 U(x)]^{-1} = \begin{pmatrix} 10.0 & 10.0 & 10.0 \end{pmatrix} \quad (2)$$

- Assuming a really simple starting point

$$x_0 = \{ \quad 1 \quad 1 \quad 1 \quad \} \quad (3)$$

- Compute the gradients and Hessian for this initial value

$$\begin{aligned} \nabla_x U(x) &= ( \quad 0.2 \quad 0.3 \quad 0.4 \quad ) \\ \nabla_{xx}^2 U(x) &= \begin{pmatrix} -0.16 & 0.06 & 0.08 \\ 0.06 & -0.21 & 0.12 \\ 0.08 & 0.12 & -0.24 \end{pmatrix} \end{aligned} \quad (4)$$

- What is the distance to the constraint?
  - To compute the distance to the constraint along the projection, let us start by computing the amount of slack to the constraint

$$c_0 = b - Ax'_0 = 94 \quad (5)$$

- How many steps along that path would it require to get to the constraint?

$$\gamma_i = \frac{c_0}{Ap} = 1.56667 \quad (6)$$

- To demonstrate the distance to the constraint, let us multiply the step by this ratio

$$x^* = x_0 + \gamma_i p_1 = \begin{pmatrix} 16.6667 & 16.6667 & 16.6667 \end{pmatrix} \quad (7)$$

$$b - Ax^{*'} = 0.0$$

- Assuming a Newton-Raphson default step ( $\alpha_1 = 1$ ) the step is then

$$\alpha_1^* = \min(\alpha_1 = 1, \gamma_i = 1.56667) \quad (8)$$

so the new point is

$$x_1 = x_0 + 1p_1 = \begin{pmatrix} 11.0 & 11.0 & 11.0 \end{pmatrix} \quad (9)$$

- Start by computing the gradient and Hessian at the new values

$$\begin{aligned}\nabla_x U(x) &= (0.1574 \quad 0.2360 \quad 0.3147) \\ \nabla_{xx}^2 U(x) &= \begin{pmatrix} -0.01144 & 0.00429 & 0.00572 \\ 0.00429 & -0.01502 & 0.00858 \\ 0.00572 & 0.00858 & -0.017166 \end{pmatrix}\end{aligned}\quad (10)$$

- The gradient is not zero. So we compute the direction of search

$$p_2 = (110.0 \quad 110.0 \quad 110.0). \quad (11)$$

- Again, consider how far it is to the constraint

$$c_1 = b - Ax'_1 = 34.0. \quad (12)$$

This implies a maximum step of

$$\gamma_i = \frac{34.0}{Ap'_2} = 0.05152. \quad (13)$$

- Again, assuming a Newton-Raphson step of  $\alpha_2 = 1$  we have

$$\begin{aligned} \alpha_2^* &= \min(\alpha_1 = 1, \gamma_i = 0.05152) = 0.05152 \\ x_2 &= \begin{pmatrix} 11.0 & 11.0 & 11.0 \end{pmatrix} + 0.05152 \times \begin{pmatrix} 110.0 & 110.0 & 110.0 \end{pmatrix} \\ &= \begin{pmatrix} 16.6667 & 16.6667 & 16.6667 \end{pmatrix} \end{aligned} \quad (14)$$

- Thus, the result is on the constraint,

$$A(x_1 + 0.05152p_2)' = 100.0 \quad (15)$$

- We need a Nullspace

$$z = \text{NullSpace}[A] = \begin{pmatrix} -3.0 & 0.0 & 1.0 \\ -2.0 & 1.0 & 0.0 \end{pmatrix}. \quad (16)$$

- The gradient and Hessian matrix at this point are

$$\begin{aligned} \nabla_x U(x) &= \begin{pmatrix} 0.15095 & 0.22643 & 0.30191 \end{pmatrix} \\ \nabla_{xx}^2 U(x) &= \begin{pmatrix} -0.007246 & 0.002717 & 0.003623 \\ 0.002717 & -0.009510 & 0.005434 \\ 0.003623 & 0.005434 & -0.010869 \end{pmatrix}. \end{aligned} \quad (17)$$

- The projected gradient is then

$$\nabla_x U(x) Z' = \begin{pmatrix} -0.15095 & -0.07548 \end{pmatrix} \quad (18)$$

- Thus, the point is not optimal – so we need to “choose the logic.” To do this we need to know the signs of the Lagrange Multipliers.

- Note that

$$A'\lambda \neq \nabla_x U(x) \quad (19)$$

because the first-order necessary conditions are not met (the projected gradients are non-zero).

- However, we need to know the sign of the Lagrange multiplier. The suggestion is to use a  $QR$  decomposition ( $A = QR$ )

$$\begin{aligned} A'Ax &= A'b \\ R'Q'QRx &= R'Q'b \\ R'Rx &= R'Q'b \quad (\text{given that } Q'Q = I) \\ Rx &= Q'b. \end{aligned} \quad (20)$$

Here we are trying to solve for a system such that

$$R\lambda = Q'\nabla_x U(x) \quad (21)$$

- Using Mathematica's QR decomposition

$$\{ Q \ R \} = \left\{ \begin{pmatrix} \frac{1}{\sqrt{14}} & \sqrt{\frac{1}{7}} & \frac{3}{\sqrt{14}} \end{pmatrix} \ \sqrt{14} \right\}. \quad (22)$$

Note that  $QR = A$ . Thus,

$$\lambda_3 = \begin{pmatrix} \frac{1}{\sqrt{14}} & \sqrt{\frac{1}{7}} & \frac{3}{\sqrt{14}} \end{pmatrix} \begin{pmatrix} 0.15095 \\ 0.22643 \\ 0.30191 \end{pmatrix} = 0.40344 \quad (23)$$

so the constraint is still binding.

- Hence, the search direction and new point become

$$\begin{aligned} p_3 &= -(\nabla_x U(x) Z) (Z' \nabla_{xx}^2 U(x) Z)^{-1} \\ x_3 &= x_2 + p_3' Z \end{aligned} \tag{24}$$

- As a first step, we compute the gradient and Hessian at  $x_3$ .
- Next, we compute for convergence

$$\nabla_x U(x) Z = \begin{pmatrix} -0.019775 & -0.016014 \end{pmatrix} \quad (25)$$

- Next, notice that the  $QR$  factorization does not change.  
Hence, the value of the Lagrange multiplier becomes

$$\lambda_4 = Q \nabla_x U(x)' = 0.42475. \quad (26)$$

- So we keep the constraints. The direction of search and new value of  $x$  becomes

$$\begin{aligned} p_4 &= -(\nabla_x U(x) Z) (Z' \nabla_{xx}^2 U(x) Z)^{-1} \\ x_4 &= x_3 + p_4' Z \end{aligned} \tag{27}$$

- At this new point, the projected gradient vanishes.

- The general objective is to maximize or minimize a nonlinear objective function subject to nonlinear constraints.
- Gill, Murray and Wright lay out the basic problems

$$\begin{array}{ll} \min_x f(x) & \\ \text{s.t. } c_i(x) = 0 \ i = 1, \dots, t & \end{array} \quad \text{and} \quad \begin{array}{ll} \min_x f(x) & \\ \text{s.t. } c_i(x) \geq 0 \ i = 1, \dots, t & \end{array} \quad (28)$$

- These optimization problems are much more difficult than the equality or inequality constraints posed in the preceding sections due to the difficulties involved in maintaining feasibility.
  - By the formulation of the null-space in the linear equality scenario, it was always possible to guarantee that  $x_{t+1}$  was feasible given that  $x_t$  was feasible.
  - Similarly, expansion to linear inequality constraints only added caveats to the step length algorithm and checks on whether a constraint could be deleted.
  - However, the solution of nonlinear constraints may be difficult, if not impossible, without the incorporation of a nonlinear objective function.
- I want to discuss three different methodologies for optimization under nonlinear constraints.
  - The first two are the use of penalty and barrier functions.
  - The third procedure is the projected augmented Lagrangian method.

- The penalty and barrier function procedures are very similar.
- The primary concept of both procedures is to append an additional term onto the objective function which imposes a cost for violating the constraint.
- Taking the Penalty function first, assume that we want to optimize

$$\begin{array}{ll} \min_x & x^2 \\ \text{s.t.} & x - 1 = 0 \end{array} \quad (29)$$

- Granted that the constraint is linear, and could be solved using a projected quasi-Newton algorithm, a quadratic penalty function for this problem could be specified as

$$P_Q(x, \rho) = x^2 + \frac{\rho}{2} (x - 1)^2 \quad (30)$$

where the  $\rho$  term yields a penalty for violating the constraint.

- Solving this problem while increasing the penalty parameter yields an optimal solution with the constraint imposed.