SySc 512, Session 11, Constrained Opt

General Constrained Optim Problem:

Find $\vec{x}^* \in \underline{X} \subset \mathbb{R}^n$ that minimizes $f(\vec{x})$

Subject to the constraints:

$c_i(\vec{x}^*) = 0$ (equality constraint)

Example 1: $f(x,y) = \frac{1}{2}(x^2+y^2)$

$c(x) = x+y-1$

Constraint slices through objective fuct

At $\vec{x}^*$: Gradients of f & c are parallel

$$\Rightarrow \nabla_x f(\vec{x})\Big|_{\vec{x}=\vec{x}^*} = \lambda_i \nabla c_i(\vec{x})\Big|_{\vec{x}=\vec{x}^*}$$

(proportionality constant)

Define "Lagrangian" to obtain opt. conditions:

$$\mathcal{L}(\vec{x},\vec{\lambda}) = f(\vec{x}) - \vec{\lambda}\cdot\vec{c}(\vec{x})$$

w/ $\vec{\lambda}$ are Lagrange multipliers

1st Order Conditions

$$\nabla_{(x,\lambda)}\mathcal{L}(\vec{x},\vec{\lambda})\Big|_{\substack{\vec{x}=\vec{x}^* \\ \vec{\lambda}=\vec{\lambda}^*}} = 0$$

$$\begin{bmatrix} \nabla_x(f(\vec{x}) - \vec{\lambda}\cdot\vec{c}(\vec{x})) \\ \vec{c}(\vec{x}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

In 2-Dim w/ 1-constraint:

$$\frac{\partial}{\partial x_1}\mathcal{L} = \frac{\partial}{\partial x_1} f(x_1,x_2) - \lambda\frac{\partial}{\partial x_1} c(x_1,x_2) = 0$$

$$\frac{\partial}{\partial x_2}\mathcal{L} = \frac{\partial}{\partial x_2} f(x_1,x_2) - \lambda\frac{\partial}{\partial x_2} c(x_1,x_2) = 0$$

$$\frac{\partial}{\partial \lambda}\mathcal{L} = c(x_1,x_2) = 0$$

(λ is like the strength of the constraint force)

SySc 512, Session 11, (cont.)

Examp 1

$$\mathcal{L}(x, y, \lambda) = \tfrac{1}{2}(x^2 + y^2) - \lambda(x + y - 1)$$

$\Rightarrow$

$$\nabla \mathcal{L} = \begin{bmatrix} x - \lambda \\ y - \lambda \\ x + y - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \left.\begin{matrix} \lambda = x \\ \lambda = y \end{matrix}\right\} \Rightarrow x = y$$

$$x = 1 - y \Rightarrow 2x = 1$$

$\Rightarrow x^* = y^* = \frac{1}{2}$, $\lambda = \frac{1}{2}$

2nd Order Conditions

Project down to accessable space:

$\Rightarrow$ perpendicular to gradient of constraint.

If $\vec{n} \cdot (\nabla C) = 0$ $\Rightarrow$ $\vec{n}$ in nullspace of ∇C.

Project Hessian of $\mathcal{L}$ to nullspace of ∇C

Find eigenvalues of:

$$H_0 = \vec{n}^T (\nabla_x^2 \mathcal{L}(\vec{x}))\vec{n}$$

If eigenvalues of $H_0|_{(x^*, \lambda^*)} > 0 \Rightarrow$ min.

Examp 1 $\nabla_x^2 \mathcal{L}(x, y, \lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$\nabla_x C = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, (\nabla C) \cdot \vec{n} = 0 \Rightarrow \vec{n} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Project: $\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2$

Since $2 > 0 \Rightarrow$ minimum

SySc 512, Session 11 (cont)

Constr Opt, Examp 2

$$\min f(\vec{x}) = \tfrac{1}{2}(x_1 - 1)^2 + \tfrac{1}{2} x_2^2$$
$$\text{s.t} \quad c(\vec{x}) = -x_1 + b\, x_2^2 = 0$$

1st Order Condition

(a) $\partial_1 \mathcal{L} = (x_1 - 1) + \lambda = 0$

(b) $\partial_2 \mathcal{L} = x_2 - 2\lambda b x_2 = x_2(1 - 2\lambda b) = 0$

(c) $\partial_\lambda \mathcal{L} = -x_1 + b x_2^2 = 0$

I (b) $\Rightarrow x_2 = 0$, $\Rightarrow$ (c) $x_1 = 0$, $\Rightarrow$ (a) $\lambda = 1$, $\Rightarrow (\vec{x}^*, \lambda^*) = (0, 0, 1)$

(b) $\Rightarrow \lambda = \frac{1}{2b}$, $\Rightarrow$ (a) $x_1 = 1 - \frac{1}{2b}$, & (c) $x_2^2 = \frac{1}{b}\left(1 - \frac{1}{2b}\right)$

$\Rightarrow (\vec{x}^*, \lambda^*) = \left(1 - \frac{1}{2b}, \pm \frac{1}{b}\sqrt{b - \frac{1}{2}}, \frac{1}{2b}\right)$

2nd Order Condition

If $\vec{n} \cdot (\nabla_x c) = 0$, & eigenvalues of $(\vec{n}^T (\nabla_x^2 \mathcal{L}) \vec{n}) > 0$

$\Rightarrow$ minimum.

$$\nabla c = \begin{bmatrix} -1 \\ 2b x_2 \end{bmatrix}, \quad \vec{n} = \begin{bmatrix} 2b x_2 \\ 1 \end{bmatrix}$$

$$[\nabla_x^2 \mathcal{L}] = \begin{bmatrix} \partial_{11}^2 \mathcal{L} & \partial_{12}^2 \mathcal{L} \\ \partial_{21}^2 \mathcal{L} & \partial_{22}^2 \mathcal{L} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 - 2\lambda b \end{bmatrix}$$

Project onto null-space:

$$\begin{bmatrix} 2b x_2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 - 2\lambda b \end{bmatrix} \begin{bmatrix} 2b x_2 \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 2b x_2 & 1 \end{bmatrix} \begin{bmatrix} 2b x_2 \\ 1 - 2\lambda b \end{bmatrix} = 4b^2 x_2^2 + 1 - 2\lambda b$$

I If $(\vec{x}^*, \lambda^*) = (0, 0, 1)$

$\Rightarrow 1 - 2b > 0$ if $b < \frac{1}{2}$ $\Rightarrow \boxed{\text{min}}$

if $b > \frac{1}{2}$ $\Rightarrow \boxed{\text{max}}$

II If $(\vec{x}^*, \lambda^*) = \left(1 - \frac{1}{2b}, \pm \frac{1}{b}\sqrt{b - \frac{1}{2}}, \frac{1}{2b}\right)$

$\Rightarrow 4\cancel{b^2}\left(\cancel{\tfrac{1}{b}}\right)^2 \left(b - \tfrac{1}{2}\right) + 1 - \frac{2b}{2b}$

$= 4\left(b - \frac{1}{2}\right) > 0$ if $b > \frac{1}{2}$ $\Rightarrow \boxed{\text{min}}$

SySc 512 Session 11, (cont.)

Rayleigh Quotient: $G(\vec{v}) = \dfrac{\vec{v}^T A \vec{v}}{\vec{v}^T \vec{v}}$, $A = A^T$

Show: ~~Maxima~~ Eigenvectors of A are minimizers of $G(\vec{v})$, if $\vec{v}^T\vec{v} = 1$

$\Rightarrow$

Minimize $\vec{v}^T A \vec{v}$
subject to $\vec{v}^T \vec{v} = 1$

(More generally, $\vec{v}$ is extremum)

Lagrangian: $\mathcal{L}(\vec{v}) = \vec{v}^T A \vec{v} + \lambda(1 - \vec{v}^T\vec{v})$

1st Order Cond: $\nabla \mathcal{L} = 0$

$$\frac{\partial}{\partial v_j}\mathcal{L} = \frac{\partial}{\partial v_j}\left(\sum_{i,k} A_{ik} v_i v_k + \lambda(1 - \sum_i v_i^2)\right)$$

$$= \sum_k A_{jk} v_k + \sum_i A_{ij} v_i - 2\lambda v_j$$

$$= 2\left(\sum A_{jk} v_k - \lambda v_j\right)$$

$\Rightarrow$

$$\nabla \mathcal{L} = 2(A\vec{v} - \lambda \vec{v}) = 0 \Rightarrow \boxed{A\vec{v} = \lambda\vec{v}}$$

Thus, $\vec{v}$ is an eigenvector & λ is eigenvalue of A.

Sensitivity to constraint

Minimize $f(x_1, x_2)$
subject to $c(\vec{x}) = g(x_1, x_2) - b$

Lagrangian: $\mathcal{L}(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda[b - g(x_1, x_2)]$

* $\longrightarrow$ What is the change in $\mathcal{L}$ caused by changing b?

$$\frac{d}{db}\mathcal{L} = \partial_1 f \frac{dx_1}{db} + \partial_2 f \frac{dx_2}{db} + [b - g(\vec{x})]\frac{d\lambda}{db}$$
$$+ \lambda\left(1 - \partial_1 g \frac{dx_1}{db} - \partial_2 g \frac{dx_2}{db}\right)$$
$$= (\partial_1 f - \lambda \partial_1 g)\frac{dx_1}{db} + (\partial_2 f - \lambda\partial_2 g)\frac{dx_2}{db} + [b - g]\frac{d\lambda}{db} + \lambda$$

At $(\vec{x}^*, \lambda^*)$, $\partial_1 f - \lambda^* \partial_1 g = 0$, $\partial_2 f - \lambda^* \partial_2 g = 0$
& $b - g(x_1^*, x_2^*) = 0$

$\Rightarrow \frac{d}{db}\mathcal{L} = \lambda^*$. Thus, λ^* gives sensitivity of optimum on b.