Lagrangian-based Methods

- **We seek**

$$\min_{x \in R^n} F(x) \text{ subject to } c(x) = 0 \in R^m$$

- **A sol'n (x^*) minimizes the *Lagrangian***

$\mathcal{L}(x, \lambda^*)$ **in the null-space** $\mathcal{N}(A(x^*))$.

Augmented Lagrangian Methods

- Using ideas of unconstrained minimization we might try to minimize

$$\mathcal{L}_A(x, \lambda^*, \rho) \equiv \mathcal{L}(x, \lambda^*) + \rho \, \|c(x)\|^2$$

- There is a finite $\bar{\rho} > 0$ such

that (x^*, λ^*) minimizes

$$\mathcal{L}_A(x, \lambda^*, \rho)$$

for all $\rho > \bar{\rho}$.

- **Also known as the *Method of Multipliers* (Hestenes)**

Projected Lagrangian Methods

- Using linearly constrained minimization we might try to

$$\min \mathcal{L}(x_k + p, \lambda^*) \;\; p \in \mathcal{N}(A(x_k))$$

- Note that

$$\mathcal{L}(x_k + p, \lambda^*) \equiv F(x_k + p) + \lambda^{*T} c(x_k + p)$$

- We need an *estimate* for λ^*. This can be done, for example, by least-squares estimate

$$\min_{\lambda} \|g_k - A_k^T \lambda\|^2$$

Projected Lagrangian Methods

- **Our initial formulation is**

$$\min_{x \in R^n} \Phi_{LC}(x) \equiv F(x) \; - \; \hat{\lambda}^T c(x)$$

- **Subject to**

$$c(\hat{x}) \; + \; A(\hat{x}) \cdot (x - \hat{x}) = 0 \in R^m$$

- $\hat{x}$ **and** $\hat{\lambda}$ **are current**

estimates of the solution of our original problem

Projected Lagrangian Methods

- To solve the (linearly constrained) sub-problem we form its Lagrangian

$$\mathcal{L}_{LC}(x, \eta) \equiv \Phi_{LC}(x) - \eta^T * [A(\hat{x}) \cdot (x - \hat{x}) + \hat{c}]$$

- **Optimality for the sub-problem leads to**

$$\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{A}^T(\hat{\boldsymbol{x}})\hat{\boldsymbol{\lambda}} = \boldsymbol{A}^T(\hat{\boldsymbol{x}})\boldsymbol{\eta}$$

- **Suppose $\hat{\boldsymbol{x}} = \boldsymbol{x}^*$ and $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^*$ is the solution to the original**

problem - then

$$\boldsymbol{g}(\boldsymbol{x}^*) - \boldsymbol{A}^T(\boldsymbol{x}^*)\boldsymbol{\lambda}^* = \boldsymbol{A}^T(\boldsymbol{x}^*)\boldsymbol{\eta}$$

$$\longrightarrow \boldsymbol{\eta} = \mathbf{0} \in \boldsymbol{R}^m$$

- **This suggests that we modify our sub-problem**

$$\Phi_{LC}(\boldsymbol{x}) \equiv \boldsymbol{F}(\boldsymbol{x}) \;-\; \hat{\boldsymbol{\lambda}}^T \boldsymbol{c}(\boldsymbol{x}) + \hat{\boldsymbol{\lambda}}^T \left[\boldsymbol{A}(\hat{\boldsymbol{x}})(\boldsymbol{x} - \hat{\boldsymbol{x}})\right]$$

- **Now the optimality condition for the modified sub-problem is**

$$\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{A}^T(\hat{\boldsymbol{x}})\hat{\boldsymbol{\lambda}} + \boldsymbol{A}^T(\hat{\boldsymbol{x}})\hat{\boldsymbol{\lambda}} = \boldsymbol{A}^T(\hat{\boldsymbol{x}})\boldsymbol{\eta}$$

- **At $\hat{\boldsymbol{x}} = \boldsymbol{x}^*$, $\hat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}^*$ the modified sub-problem has solution $\boldsymbol{x} = \boldsymbol{x}^*$, $\boldsymbol{\eta} = \boldsymbol{\lambda}^*$.**

Projected Lagrangian Methods

- **The sub-problem is *linearly constrained*, but the cost function Φ_{LC} is general.**
- **Suppose we use a *quadratic approximation* for Φ_{LC}.**

- **This is conveniently done in terms of**

$$p \equiv x - \hat{x}$$

- **Expand the cost function to**

second-order

$$\Phi_{LC}(\hat{x}+p) \approx \Phi_{LC}(\hat{x}) + \left[\frac{\partial \Phi_{LC}}{\partial x}\right]_{\hat{x}}^{T} p + (1/2)p^{T}\left[\frac{\partial^2 \Phi_{LC}}{\partial x^2}\right]_{\hat{x}}^{T} p$$

➤ **This leads to**

$$\Phi_{LC}(\hat{x}+p) \approx \Phi_{LC}(\hat{x}) + g(\hat{x})^T p + (1/2) p^T \left[\frac{\partial^2 \mathcal{L}}{\partial x^2}\right]^T_{(\hat{x},\hat{\lambda})} p$$

Sequential Quadratic Programming

- Second-order ‘cost’ function (Φ_{PL}) and first-order constraints

- **Leads to the sub-problem**

$$\min_{p \in R^n} g_k^T p + (1/2) p^T H_k p$$

$$\text{subject to} \quad A_k p + c_k = 0$$

- **Optimality for the sub-problem**

leads to:

$$\begin{bmatrix} \boldsymbol{H_k} & -\boldsymbol{A_k^T} \\ \boldsymbol{A_k} & \boldsymbol{0} \end{bmatrix} \begin{pmatrix} \boldsymbol{p_k} \\ \boldsymbol{\eta_k} \end{pmatrix} = \begin{pmatrix} -\boldsymbol{g_k} \\ -\boldsymbol{c_k} \end{pmatrix}$$

- **Note g_k is the gradient of F and H_k is the estimate of the Hessian of the Lagrangian**
- **Related linear system arises**

from Newton's method for the optimality system

$$\begin{aligned} \boldsymbol{g}(\boldsymbol{x}) \; -\boldsymbol{A}(\boldsymbol{x})^T\boldsymbol{\lambda} &= \boldsymbol{0} \\ \boldsymbol{c}(\boldsymbol{x}) &= \boldsymbol{0} \end{aligned}$$

Projected Lagrangian Example

- The problem is to

$$\min_{x \in R^2} x_1 x_2^2$$

- subject to

$$2 - x_1^2 - x_2^2 = 0$$

- **We find that**

$$\boldsymbol{g(x) = \begin{bmatrix} x_2^2 \\ 2\, x_1\, x_2 \end{bmatrix}}$$

and

$$\boldsymbol{A(x) = \begin{bmatrix} -2\, x_1 & -2\, x_2 \end{bmatrix}}$$

- **Further calculations reveal that**

$$\left[\frac{\partial^2 \mathcal{L}}{\partial x^2}\right] = \begin{bmatrix} 2\,\lambda & 2\,x_2 \\ 2\,x_2 & 2\,(\lambda + x_1) \end{bmatrix}$$

- **A solution to the problem is**

$$x^* = \begin{bmatrix} -\sqrt{(2/3)} \\ -\sqrt{(4/3)} \end{bmatrix} \approx \begin{bmatrix} -0.8165 \\ -1.1547 \end{bmatrix}$$

$$\lambda^* = \sqrt{(2/3)} \approx 0.8165$$

Projected Lagrangian Example

- **We begin with the estimate**

$$\hat{\boldsymbol{x}}_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \text{and} \quad \hat{\boldsymbol{\lambda}}_0 = 1$$

- **This leads to**

$$g(\hat{x}_0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \;\; A(\hat{x}_0) = \begin{bmatrix} 2 & 2 \end{bmatrix},$$

$$\hat{c}_0 = 0, \;\; \left[\frac{\partial^2 \mathcal{L}}{\partial x^2}\right]_0 = \begin{bmatrix} 2 & -2 \\ -2 & 0 \end{bmatrix}$$

- **The quadratic sub-problem is**

$$\min_{p \in R^2} \begin{bmatrix} 2 & 2 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

$$+(1/2)[p_1 \; p_2] \begin{bmatrix} 2 & -2 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

- **subject to the *linear constraint***

$$2\,p_1 + 2\,p_2 + 0 = 0$$

- **The solution is**

$$p = \begin{bmatrix} 1/6 \\ -1/6 \end{bmatrix} \text{ and } \eta = 5/6$$

Projected Lagrangian Example

- The next iterate is then

$$\hat{\boldsymbol{x}}_1 = \begin{bmatrix} -5/6 \\ -7/6 \end{bmatrix} \quad \text{and} \quad \hat{\boldsymbol{\lambda}}_1 = 5/6$$

- The subsequent quadratic

sub-problem has

$$\boldsymbol{g(\hat{x}_1) = \begin{bmatrix} 49/36 \\ 70/36 \end{bmatrix}, \ \hat{c}_1 = -1/18,}$$

$$\boldsymbol{A(\hat{x}_1) = \begin{bmatrix} 10/6 & 14/6 \end{bmatrix},}$$

$$\left[\frac{\partial^2 \mathcal{L}}{\partial x^2}\right]_1 = \begin{bmatrix} 10/6 & -14/6 \\ -14/6 & 0 \end{bmatrix}$$

- **The solution is**

$$p = \begin{bmatrix} 1/60 \\ 1/84 \end{bmatrix} \text{ and } \eta = 49/60$$

➤ Leading to the next iterate

$$\hat{\boldsymbol{x}}_2 = \begin{bmatrix} -49/60 \\ -97/84 \end{bmatrix} \approx \begin{bmatrix} -.8167 \\ -1.1548 \end{bmatrix}$$

$$\text{and} \quad \hat{\boldsymbol{\lambda}}_1 = 49/60 \approx .8167$$