Linear Equality Constraints

- **Minimize a smooth function of several variables $F(x)$, $x \in \mathbb{R}^n$ subject to linear equality constraints**

$$Ax = b \in \mathbb{R}^m, \quad m < n$$

Constraint Set

- We can complet**el**y characterize the feasible points as

$$\{x = x_1 + p | \text{where } p \in \mathcal{N}(A)\}$$

and x_1 is some solution of $Ax = b$.

- **Let the columns matrix Z be a representation for $\mathcal{N}(A)$. Z will have $nc = n - \text{rank}(A)$ columns.**
- **We write $x = x_1 + Zp_z$ where $p_z \in \mathbb{R}^{nc}$ is arbitrary.**
- **If we define**

$$\hat{F}(p_z) \equiv F(x_1 + Zp_z)$$

we can minimize $\hat{F}$ as an *unconstrained* problem.

➤ Note

$$\nabla_Z \hat{F} = Z^T \left(\nabla F\right)$$

and

$$\nabla_Z^2 \hat{F} = Z^T \left(\nabla^2 F\right) Z$$

Null Space Representation

- **Our usual procedure uses a Q-R decomposition to get $Z = \mathcal{N}(A)$**
- **In many problems we can split the unkown x into two sub-vectors x_V and x_U and**

compatibly partition the A matrix to write

$$\boldsymbol{Ax} = [\boldsymbol{VU}] \begin{bmatrix} \boldsymbol{x_V} \\ \boldsymbol{x_U} \end{bmatrix} = \boldsymbol{b}.$$

- **The x_U sub-vector is interpreted as the *independent* variables, while the x_V**

sub-vector is interpreted as the *dependent* variables.

$$\boldsymbol{A x} = \boldsymbol{b} \rightarrow \boldsymbol{x_V} = \boldsymbol{V}^{-1}\left[\boldsymbol{b} - \boldsymbol{U x_U}\right]$$

- **This leads to a null-space representation**

$$\begin{bmatrix} -\boldsymbol{V}^{-1}\boldsymbol{U} \\ \boldsymbol{I} \end{bmatrix}.$$

Quadratic Programming (QP)

➤ **An important special case of linearly constrained problems occurs when F is a quadratic function**

- **We seek to minimize**

$$(1/2)x^T G x + c^T x$$

subject to

$$Ax = b$$

- **As noted for the general problem with linear equality**

constraints we write

$$x = \bar{x} + p,$$

where $\bar{x}$ is some feasible point ($A\bar{x} = b$) and $p \in \mathcal{N}(A)$.

- **Using this representation we get the equivalent problem**

$$\min_{p} \; (1/2)p^T G p + (G\bar{x} + c)^T p$$

subject to

$$\boldsymbol{Ap = 0}$$

- **Since**

$$\boldsymbol{p \in \mathcal{N}(A) \rightarrow p = Zp_z}$$

we get the unconstrained

problem

$$\min_{\boldsymbol{p_z} \in \mathbb{R}^{n-m}} (1/2)\boldsymbol{p_z}^T \boldsymbol{Z}^T \boldsymbol{G} \boldsymbol{Z} \boldsymbol{p_z} + (\boldsymbol{G}\bar{\boldsymbol{x}} + \boldsymbol{c})^T \boldsymbol{Z} \boldsymbol{p_z}$$

Lagrange Mulitplier Estimates

- At a critical point $\nabla \mathcal{L} = 0$, or

$$\nabla F(x^*) = A^T \lambda \quad (*)$$

- At a typical x the system (*) is an incompatible, system (n equations in m unknowns).

- **For such points ($\tilde{x}$) we might replace (*) with**

$$\min_{\lambda \in \mathbb{R}^m} \|A^T\lambda - \nabla F(\tilde{x})\|^2$$

- **These are sometimes called *projection* mulitpliers.**

- **An alternative is to use the variable-splitting idea discussed**

above and solve the $\boldsymbol{m}$ by $\boldsymbol{m}$ system

$$\boldsymbol{V}^{\boldsymbol{T}}\boldsymbol{\lambda} = \boldsymbol{\nabla}_{\boldsymbol{x}_{\boldsymbol{V}}}\boldsymbol{F}$$

Inequality Constraints
Karush-Kuhn-Tucker

- **If we consider *inequality* constraints**

$$\boldsymbol{Ax \geq b} \quad \text{or} \quad \boldsymbol{Ax - b \geq 0}$$

we get as a necessary condition

for a minimizer

$$\nabla F(x^*) = A^T \lambda$$

and

$$\lambda \geq 0$$

- ***inactive* constraints correspond to rows of A with**

$$A_j x > b_j$$

Such constraints will have $\boldsymbol{\lambda}_{\jmath} = 0$