2c

NEWTON TYPE ALGORITHMS IN OPTIMIZATION

EQUALITY CONSTRAINED OPTIMIZATION

Equality constraints only, first

$f: \mathbb{R}^n \to \mathbb{R}$

$g: \mathbb{R}^n \to \mathbb{R}^n$

(Both assumed to be smooth)

$$\min_{w \in \mathbb{R}^N} f(w) \quad \text{s.t.} \quad g(w) = 0$$

WE CAN USE NEWTON'S TYPE METHODS, OR SOME VARIANT OF IT, TO SOLVE THE KKT CONDITIONS (which are nonlinear)

$$\begin{cases} \nabla_w \mathcal{L}(w,\lambda) = 0 \\ g(w) = 0 \end{cases} ; \text{ from } \mathcal{L}(w,\lambda,\mu) = f(w) - \lambda^T g(w) - \underbrace{\mu^T h(w)}_{=0}$$

(THE LAGRANGIAN FUNCTION)

THIS IS NOW A ROOT FINDING PROBLEM

WE SIMPLIFY THE NOTATION BY DEFINING THE FOLLOWING TERMS

$$z = \begin{bmatrix} w \\ \lambda \end{bmatrix} \quad \text{AND} \quad R(z) = \begin{bmatrix} \nabla_w \mathcal{L}(w,\lambda) \\ g(w) \end{bmatrix},$$

SO THAT $z \in \mathbb{R}^{N+N_g}$ AND $R: \mathbb{R}^{N+N_g} \longrightarrow \mathbb{R}^{N+N_g}$

WE CAN WRITE THE ROOT-FINDING PROBLEM COMPACTLY AS

$$R(z) = 0$$

CONSIDER AN INITIAL SOLUTION z_0, THE NEWTON'S METHOD GENERATES A SEQUENCE OF ITERATES $(z_k)_{k=0}^{\infty}$ BY LINEARIZING THE NONLINEAR EQUATION AT THE CURRENT ITERATION POINT

$$R(z_k) + \left.\frac{\partial R}{\partial z}\right|_{z_k} (z - z_k) = 0$$

~~AND~~ THE NEXT ITERATE IS THE SOLUTION TO THE LINEARISED EQ.

$$z_{k+1} = z_k - \left[\left.\frac{\partial R}{\partial z}\right|_{z_k}\right]^{-1} R(z_k)$$

IN TERMS OF GRADIENTS

$$\nabla_w \mathcal{L}(w_k,\lambda_k) + \underbrace{\nabla^2_w \mathcal{L}(w,\lambda)(w - w_k)}_{\substack{\text{LINEARISATION} \\ \text{WRT } w}} - \underbrace{\nabla g(w_k)(\lambda - \lambda_k)}_{\substack{\text{LINEARISATION} \\ \text{WRT } \lambda}} = 0$$

By definition :

$$\begin{cases} \lambda_{k+1} = \lambda_k + \Delta\lambda_k \\ \nabla \mathcal{L}(w_k,\lambda_k) = \nabla f(w_k) - \nabla g(w_k)\lambda \end{cases}$$

To calculate $\nabla\lambda_k$ and ∇w_k is equivalent to solving

$$\begin{bmatrix} \nabla f(w_k) \\ g(w_k) \end{bmatrix} + \begin{bmatrix} \nabla^2 \mathcal{L} & \nabla g \\ \nabla g^T & 0 \end{bmatrix} \begin{bmatrix} \Delta w_k \\ -\lambda_{k+1} \end{bmatrix} = 0$$

FOR EQUALITY CONSTRAINED OPTIMIZATION PROBLEMS, THE LINEAR SYSTEM $R(z_k) + \left.\frac{\partial R}{\partial z}\right|_{z_k} (z - z_k) = 0$ HAS A SPECIFIC FORM

RESIDUALS

MATRIX OF SECOND DERIVATIVES

$$\begin{bmatrix} \nabla_w \mathcal{L}(w_k, \lambda_k) \\ g(w_k) \end{bmatrix} + \underbrace{\begin{bmatrix} \nabla_w^2 \mathcal{L}(w_k, \lambda_k) & \nabla g(w_k) \\ \nabla g(w_k)^T & 0 \end{bmatrix}}_{\text{KKT matrix}} \begin{bmatrix} w - w_k \\ \lambda - \lambda_k \end{bmatrix} = 0$$

BY USING THE DEFINITION $\nabla_w \mathcal{L}(w_k, \lambda_k) = \nabla f(w_k) + \nabla g(w_k)^T \lambda_k$

WE CAN REWRITE THE SYSTEM AS

$$\begin{bmatrix} \nabla f(w_k) \\ g(w_k) \end{bmatrix} + \begin{bmatrix} \underline{\nabla_w^2 \mathcal{L}(w_k, \lambda_k)} & \nabla g(w_k) \\ \nabla g(w_k)^T & 0 \end{bmatrix} \begin{bmatrix} w - w_k \\ \lambda \end{bmatrix} = 0$$

There is no dependence on λ_k

NEWTON-TYPE
vs
EXACT NEWTON METHODS

The only dependence on λ_k is via the HESSIAN matrix (WHICH CAN BE APPROXIMATED)

Let $B_k = \nabla_w^2 \mathcal{L}(w_k, \lambda_k)$, we have the NEWTON-TYPE ITERATION

$$\begin{bmatrix} w_{k+1} \\ \lambda_{k+1} \end{bmatrix} = \begin{bmatrix} w_k \\ 0 \end{bmatrix} - \underbrace{\begin{bmatrix} B_k & \nabla g(w_k) \\ \nabla g^T(w_k) & 0 \end{bmatrix}^{-1}} \begin{bmatrix} \nabla f(w_k) \\ g(w_k) \end{bmatrix}$$

Invertibility is required
(LICQ) + (SECOND ORDER SUFFICIENT OPTIMALITY CONDITION)
+ (STRICT COMPLEMENTARITY)

No12

INEQUALITY CONSTRAINED OPTIMIZATION

minimise $f(w)$
$w \in \mathbb{R}^N$
subject to $g(w) = 0$
$h(w) \geq 0$

$f: \mathbb{R}^N \to \mathbb{R}$
$g: \mathbb{R}^N \to \mathbb{R}^{N_g}$
$h: \mathbb{R}^N \to \mathbb{R}^{N_h}$

All assumed to be smooth

TO SOLVE A GENERAL NONLINEAR OPTIMISATION PROBLEM WITH INEQUALITY CONSTRAINTS THERE EXIST TWO MAIN APPROACHES

- (NONLINEAR) INTERIOR POINT METHODS
- SEQUENTIAL QUADRATIC PROGRAMMING

BOTH APPROACHES AIM AT SOLVING THE KKT CONDITIONS

$$
\begin{aligned}
\nabla f(w^*) + \nabla g(w^*)\lambda^* + \nabla h(w^*)\mu^* &= 0 \\
g(w^*) &= 0 \\
h(w^*) &\leq 0 \\
\mu^* &\geq 0 \\
\mu_i^* h_i(w^*) &= 0 \quad (i = 1, 2, \_\_, N_h)
\end{aligned}
$$

NONLINEAR INTERIOR POINT METHODS REPLACE THE KINKINESS OF THE NON-SMOOTH COMPLEMENTARITY CONDITIONS WITH A SMOOTH APPROXIMATION

$\rightsquigarrow$ THE L SHAPED SET IS REPLACED BY A HYPERBOLA (typically)

TO GET

$$
\begin{aligned}
\nabla f(w^*) + \nabla g(w^*)\lambda^* + \nabla h(w^*)\mu^* &= 0 \\
g(w^*) &= 0 \\
\underbrace{\mu_i^* h_i(w^*) + \tau}_{} &= 0
\end{aligned}
$$

SO THAT $-h_i(w^*)$ AND μ_i^* ARE BOTH POSITIVE AND ON AN HYPERBOLA FOR τ SMALL

SEQUENTIAL QUADRATIC PROGRAMMING AT EACH ITERATION IT SOLVES AN INEQUALITY CONSTRAINED QUADRATIC PROGRAM OBTAINED BY LINEARISING THE OBJECTIVE FUNCTION AND THE CONSTRAINT FUNCTIONS

$$\underset{w \in \mathbb{R}^n}{\text{minimise}} \quad \nabla f(w)(w - w_k) + \tfrac{1}{2}(w - w_k)^T B (w - w_k)$$

$$\text{subject to} \quad g(w_k) + \nabla g^T(w_k)(w - w_k) = 0$$

$$h(w_k) + \nabla h^T(w_k)(w - w_k) \leq 0$$