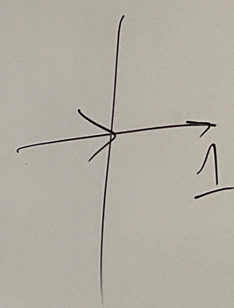
How to test for Feasibility

$x^{(n)} \in \Omega$. Based on the problem itself. Is the parameter correct? Is the probability between 0 & 1? Is Σ pos def?

$\theta = (\mu, \sigma)$

$N\left(\begin{pmatrix}\mu_1 \\ \vdots \\ \mu_p\end{pmatrix}, \begin{pmatrix}\sigma_{11} & & \\ & \ddots & \\ & & \sigma_{pp}\end{pmatrix}\right)$

$\underbrace{\mu_1, \ldots, \mu_p}_{p \text{ length}}, \underbrace{\sigma_{11}, \ldots, \sigma_{pp}}_{p(p+1)/2 \text{ length}}$



① Newton's Method

a) Compute direction (∇, ∇^2)

b) Compute $\theta^{(new)} = \theta^{(old)} + d$

c) Grab Σ from θ & check if Σ is pos. def. (feasibility)

→ check if all e-values > 0

Stopping Conditions

MRE, grad_norm

Hwk 3.3 & 3.4

② Steepest Ascent

a) Compute direction (∇ only)

b), c) same as Newton

feasibility is same as Newton

* we use Step halving if it is feasible but too much

Stopping Conditions

MRE & grad_norm

Hwk 3.2

③ Secant Method

a) Start w/ 2 points

b) Compute ∇ w.r.t. the 2 pts.

c) Use above to compute next pt. w/ direction

No feasibility present.

Stopping Conditions

mod_rel_error & grad_norm

Hwk 2.2

④ Fisher-Scoring

a) Compute direction w/ $\nabla, (E(-\nabla^2))^{-1}$

b, c) same as Newton

* Feasibility check is same as Newton

Stopping Conditions

mod_rel_error

&

gradient_norm

Hwk 3.3

BFGS

⑤ Quasi-Newton

- we start w/ $\theta^{(0)}$ & $B^{(n)}$ which (pos. def.) is an approx. to $-(\nabla^2)^{-1}$ of $f(x)$
- direction here is

$$B^{(n)} \cdot \nabla f(x^{(n)}) = d^{(n)}$$

- update θ by:

$$\theta^{(n+1)} = \theta^{(n)} + \alpha^{(n)} d^{(n)}$$

(use optim() in R)

Hwk 3.4

⑥ Weighted Nonlinear Least Squares (Gauss Newton)

a) look at data to see what distribution to use

b) Calculate Jacobian $= J$

c) Calculate direction as:

$$(J^T I J)^{-1} J \cdot I(y - f(x, \theta))$$

response ↑ Expected Value ↑

Stopping Conditions:

mod_rel_error

Hwk 4.2

⑦ IRWLS / linear IRLS

Reweighted

a) Gather $E(Y)$ & $Var(Y)$

$y_i = \mu_i(\theta) + \varepsilon_i = f(x_i, \theta) + \varepsilon_i$

$Var(y_i) = \sigma_i^2$

$\text{Weights} = 1/Var(y_i) = 1/\sigma_i^2$

$$(J^T \cdot W \cdot J)^{-1} J \cdot W(y - f(x, \theta))$$

for obtaining MLEs

$\beta_0 + \beta_1 x + \beta_2 x^2$ ← linear wrt β

$\beta_0 x^0 + \beta_1 x^1 + \beta_2 x^2$

$\begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} \begin{pmatrix} 1 & x & x^2 \end{pmatrix}$

$\log(\beta_0) + e^{\beta_1} x_1 + \beta_2^2$

fits nonlinear regression model with fixed weights to stabilize variance

Ratio of Convergence

Linear Convergence: $$\frac{\|x^{(n+1)} - x^*\|}{\|x^{(n)} - x^*\|} \leq k \quad , \quad k < 1$$

If this converges to 0, we get at least superlinear

If this converges to (0,1) we get linear

Quadratic Convergence: $$\frac{\|x^{(n+1)} - x^*\|}{\|x^{(n)} - x^*\|^2} < k \quad , \quad k \text{ is a constant}$$

If this converges to any constant, we get quadratic convergence

To see which method is best, check for

Global Convergence & Quadratic convergence

$$\frac{\|x^{(n+1)} - x^*\|}{\left(\|x^{(n)} - x^*\|\right)^n} \leq k \quad , \quad k < 1 \quad , \quad 1 < n < 2$$

This is superlinear if this converges

How to test if a direction is Ascending

Let d = some direction

If you do $d \cdot \nabla f(x)^T$ & it is positive, then d is ascending

Do we need Step halving?

Let $\theta^{(n)}$ be parameters at nth iteration & ℓ = likelihood

Let $\theta^{(n+1)} = \theta^{(n)}$ + direction

If $\ell(\theta^{(n+1)}) < \ell(\theta^{(n)})$, then we need step halving

Homework 4.1 uses nonlinear regression (nls function built into R that uses gauss newton) to estimate MLEs

Homework 4.2 uses IRWLS to estimate MLEs

Why don't we use step halving in IRWLS but we do in Fisher Scoring?
-Because Fisher scoring is an estimation of the log likelihood which you get close to the MLE. But in IRWLS we can't use step halving since the weights change. Also, because we are not using the log likelihood in IRWLS but the model given in the problem.