Chapter-6

Tuesday, July 6, 2021 11:51 AM

The operator $J(x): T_x M \to T_x M: \eta \mapsto \nabla_\eta \xi$ is called the Newton equation and its solution $\eta_k \in T_{x_k} M$ is called the Newton vector.

for $k = 0, 1, 2, \dots$ do
  // Solve the Newton equation where $J(x_k)\eta_k = \nabla_{\eta_k} \xi$.
$$J(x_k)\,\eta_k = -\xi_{x_k}.$$
  // Set
$$x_{k+1} = R_{x_k}(\eta_k)$$
end for

If ∇ is the Riemannian connection, then in compatibility with Riemannian metric,

$$D\langle \xi, \xi \rangle (x_k)[\eta_k] = \langle \nabla_{\eta_k} \xi, \xi \rangle + \langle \xi, \nabla_{\eta_k} \xi \rangle = -2\langle \xi, \xi \rangle_{x_k} < 0.$$

Riemannian Newton Method.

For the case $\xi = \operatorname{grad} f$,

$$\operatorname{Hess} f(x_k)\,\eta_k = -\operatorname{grad} f(x_k).$$

This gives the Riemannian Newton method which is a specific case of the previous algorithm,

for $k = 0, 1, 2, \dots$ do
  Solve the Newton Equation
$$\operatorname{Hess} f(x_k)\,\eta_k = -\operatorname{grad} f(x_k).$$
  Set
$$x_{k+1} = R_{x_k}(\eta_k).$$
endfor

In general η_k is not necessarily a descent direction. We have,

$$Df(x_k)[\eta_k] = \langle \operatorname{grad} f(x_k), \eta_k \rangle = -\langle \operatorname{grad} f(x_k), (\operatorname{Hess} f(x_k))^{-1} \operatorname{grad} f(x_k) \rangle.$$

When ∇ is a symmetric affine connection, $\operatorname{Hess} f(x_k)$ is positive-definite iff all its eigenvalues are strictly positive.

Practical methods make use of quasi-Newton updates,

$$(\operatorname{Hess} f(x_k) + E_k)\,\eta_k = -\operatorname{grad} f(x_k).$$

Rayleigh Quotient Algorithms.

Sphere -

We have the cost function f on the unit sphere,

$$f: S^{n-1} \to \mathbb{R}: x \mapsto x^T A x.$$

$$\operatorname{grad} f(x) = 2P_x(Ax) = 2(Ax - xx^T Ax).$$

where $P_x(Ax)$ is the orthogonal projector onto $T_x S^{n-1}$,

$$P_x z = z - xx^T z.$$

We pick the Riemannian connection as our affine connection,

$$\nabla_\eta \xi = P_x(D\xi(x)[\eta]).$$

Now we can apply the Newton method to the vector field $\xi = \operatorname{grad} f$,

$$\begin{aligned}\nabla_\eta \operatorname{grad} f &= 2P_x(D\operatorname{grad} f(x)[\eta]). \\ &= 2P_x(A\eta - \eta x^T Ax) \\ &= 2(P_x A P_x \eta - \eta x^T Ax).\end{aligned}$$

Thus the Newton eq. reads,

$$\begin{cases} P_x A P_x \eta - \eta x^T Ax = -P_x Ax \\ x^T \eta = 0. \end{cases}$$

Following its solution, we set $x_{k+1} = R_{x_k} \eta_k$ and make the Newton update

Grassman manifold -

Consider the cost function,

$$f: \operatorname{Grass}(p,n) \to \mathbb{R}: \operatorname{span}(Y) \mapsto \operatorname{tr}((Y^T Y)^{-1} Y^T A Y).$$

$\operatorname{Grass}(p,n)$ is a Riemannian quotient manifold,

horizontal distribution, $\mathcal{H}_Y = \{ Z \in \mathbb{R}^{n \times p} : Y^T Z = 0 \}$.

projection, $P_Y^h = (I - Y(Y^T Y)^{-1} Y^T)$.

and gradient, $\overline{\operatorname{grad} f}_Y = 2P_Y^h AY = 2(AY - Y(Y^T Y)^{-1} Y^T AY)$.

Furthermore, $\overline{\nabla_\eta \xi} = P_Y^h(D\bar{\xi}(Y)[\bar{\eta}_Y])$.

This yields the following expression,

$$\overline{\nabla_\eta \operatorname{grad} f} = P_Y^h(D\overline{\operatorname{grad} f}(Y)[\bar{\eta}_Y]) = 2P_Y^h(A\bar{\eta}_Y - \bar{\eta}_Y(Y^T Y)^{-1} Y^T AY).$$

Taking the horizontal lift of the Newton eq $\nabla_\eta \operatorname{grad} f = -\operatorname{grad} f(x)$.

$$P_Y^h(A\bar{\eta}_Y - \bar{\eta}_Y(Y^T Y)^{-1} Y^T AY) = -P_Y^h AY$$

Thus, the Newton eq. becomes,

$$\begin{cases} P_{Y_k}^h(AZ_k - Z_k(Y_k^T Y_k)^{-1} Y_k^T AY_k) = -P_{Y_k}^h(AY_k). \\ Y_k^T Z_k = 0. \end{cases}$$

x ——————— x ——————— x