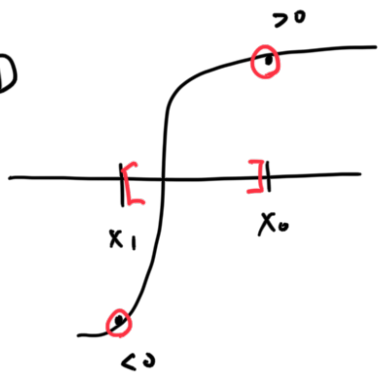
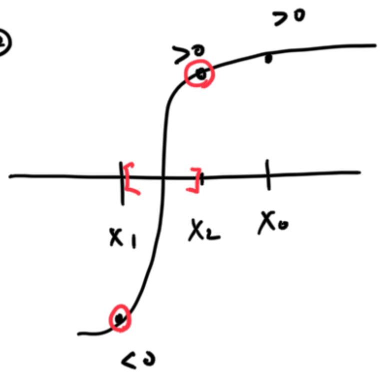
Sec. 2.1 Bisection

①

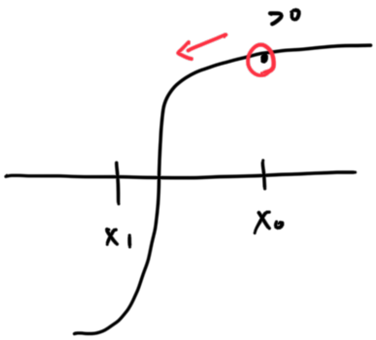


②

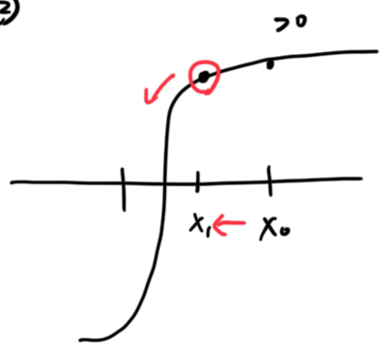


Bisection 可看成

①'

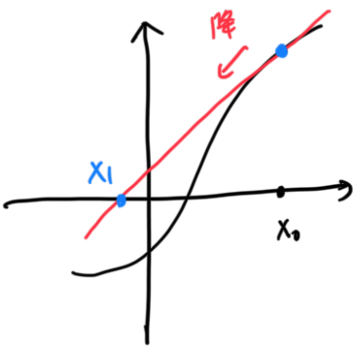


②'

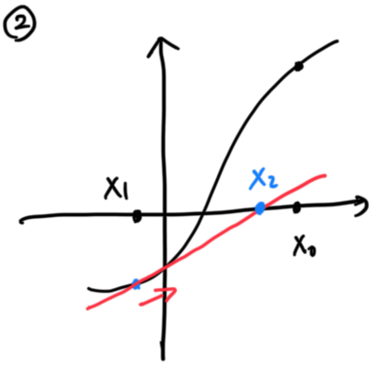


Section Newton

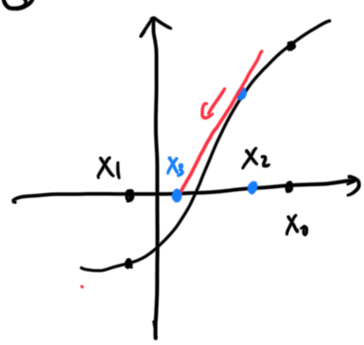
①



②



③



Section 10.4 Steepest Descent Techniques
降

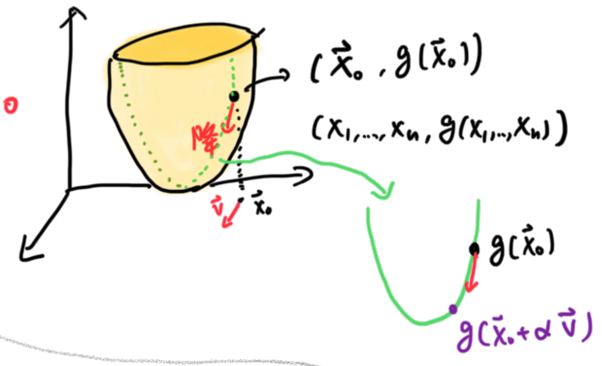
sol. $\vec{F}(\vec{x}) = \vec{0}$

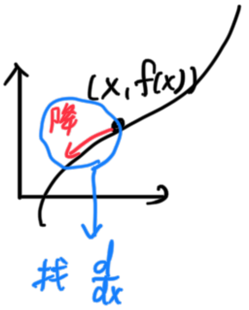
① let $\boxed{g(\vec{x}) = \sum_{i} f_i^2(\vec{x})} \geq 0$

② 找 $g(\vec{x}_0) = g_0$

③ 找 $\vec{v}$ s.t. $g(\vec{x}_0 + \vec{v})$ 降

repeat ② ③





Def.

① gradient of g at $\vec{x}$, $\nabla g(\vec{x}) = \left(\frac{\partial g}{\partial x_1}(\vec{x}), \dots, \frac{\partial g}{\partial x_n}(\vec{x})\right)^T$

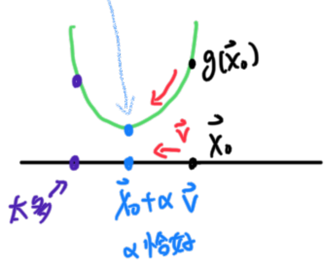
② directional derivative of g at $\vec{x}$ in the direction $\vec{v}$

slope $D_{\vec{v}}\, g(\vec{x}) = \lim_{h \to 0} \frac{g(\vec{x} + h\vec{v}) - g(\vec{x})}{h} = \vec{v}^T \cdot \nabla g(\vec{x})$

(i) 找 unit vector $\vec{v}$ s.t. $\| D_{\vec{v}}\, g(\vec{x}) \|$ max → $\vec{v} \parallel \nabla g(\vec{x})$, pick $\boxed{-\nabla g(\vec{x})}$

(ii) find α s.t. $g(\vec{x} + \alpha\vec{v}) < g(\vec{x})$

↑小夠多



Step 1. $P(x)$: U , interpolation $P_2(x)$: deg 2

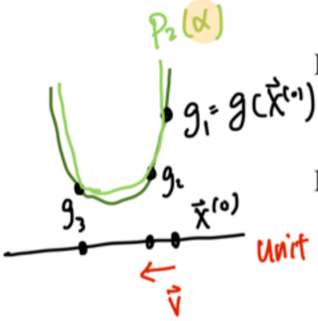
step 2. find min of $P_2(x)$

Example 1 Use the Steepest Descent method with $\mathbf{x}^{(0)} = (0, 0, 0)^t$ to find a reasonable starting approximation to the solution of the nonlinear system

$$\vec{F}(\vec{x}): \begin{bmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \\ f_3(x_1, x_2, x_3) \end{bmatrix} \begin{aligned} &= 3x_1 - \cos(x_2x_3) - \frac{1}{2} = 0, \\ &= x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 = 0, \\ &= e^{-x_1x_2} + 20x_3 + \frac{10\pi - 3}{3} = 0. \end{aligned}$$

Solution Let $g(x_1, x_2, x_3) = [f_1(x_1, x_2, x_3)]^2 + [f_2(x_1, x_2, x_3)]^2 + [f_3(x_1, x_2, x_3)]^2$. Then

$$\nabla g(x_1, x_2, x_3) \equiv \nabla g(\mathbf{x}) = \left(2f_1(\mathbf{x})\frac{\partial f_1}{\partial x_1}(\mathbf{x}) + 2f_2(\mathbf{x})\frac{\partial f_2}{\partial x_1}(\mathbf{x}) + 2f_3(\mathbf{x})\frac{\partial f_3}{\partial x_1}(\mathbf{x}),\right.$$
$$2f_1(\mathbf{x})\frac{\partial f_1}{\partial x_2}(\mathbf{x}) + 2f_2(\mathbf{x})\frac{\partial f_2}{\partial x_2}(\mathbf{x}) + 2f_3(\mathbf{x})\frac{\partial f_3}{\partial x_2}(\mathbf{x}),$$
$$\left.2f_1(\mathbf{x})\frac{\partial f_1}{\partial x_3}(\mathbf{x}) + 2f_2(\mathbf{x})\frac{\partial f_2}{\partial x_3}(\mathbf{x}) + 2f_3(\mathbf{x})\frac{\partial f_3}{\partial x_3}(\mathbf{x})\right)$$
$$= 2\mathbf{J}(\mathbf{x})^t\mathbf{F}(\mathbf{x}).$$



For $\mathbf{x}^{(0)} = (0, 0, 0)^t$, we have

$$g\left(\mathbf{x}^{(0)}\right) = 111.975 \quad \text{and} \quad z_0 = ||\nabla g\left(\mathbf{x}^{(0)}\right)||_2 = 419.554.$$

Let

$$\text{unit vector } \vec{v} = \mathbf{z} = \frac{1}{z_0}\nabla g\left(\mathbf{x}^{(0)}\right) = (-0.0214514, -0.0193062, 0.999583)^t.$$

With $\alpha_1 = 0$, we have $g_1 = g\left(\mathbf{x}^{(0)} - \alpha_1\mathbf{z}\right) = g\left(\mathbf{x}^{(0)}\right) = 111.975$. We arbitrarily let $\alpha_3 = 1$ so that

$$g_3 = g\left(\mathbf{x}^{(0)} - \alpha_3\mathbf{z}\right) = 93.5649.$$

Because $g_3 < g_1$, we accept α_3 and set $\alpha_2 = \alpha_3/2 = 0.5$. Thus

$$g_2 = g\left(\mathbf{x}^{(0)} - \alpha_2\mathbf{z}\right) = 2.53557.$$

We now find the quadratic polynomial that interpolates the data (0, 111.975), (1, 93.5649), and (0.5, 2.53557). It is most convenient to use Newton's forward divided-difference interpolating polynomial for this purpose, which has the form 3-3

$$P(\alpha) = g_1 + h_1\alpha + h_3\alpha(\alpha - \alpha_2).$$

This interpolates

$$g\left(\mathbf{x}^{(0)} - \alpha\nabla g\left(\mathbf{x}^{(0)}\right)\right) = g\left(\mathbf{x}^{(0)} - \alpha\mathbf{z}\right)$$

$P(\alpha) = g_1 + h_1\alpha + h_3\alpha(\alpha - \alpha_2)$

$P'(\alpha) = h_1 + h_3(\alpha - \alpha_2) + h_3\alpha$

$= 2h_3\alpha + (h_1 - \alpha_2 h_3)$

$\therefore$ roots: $\boxed{\dfrac{\alpha_2 h_3 - h_1}{2h_3} = \alpha_0}$

① $\|\vec{X}^{(k)} - \vec{X}^{(k-1)}\|_\infty < TOL$

② $g(\vec{X}^{(k)}) < TOL$

We now find the quadratic polynomial that interpolates the data $(0, 111.975)$, $(1, 93.5649)$, and $(0.5, 2.53557)$. It is most convenient to use Newton's forward divided-difference interpolating polynomial for this purpose, which has the form

3-3

$$P(\alpha) = g_1 + h_1\alpha + h_3\alpha(\alpha - \alpha_2).$$

This interpolates

$$g\left(\mathbf{x}^{(0)} - \alpha\nabla g(\mathbf{x}^{(0)})\right) = g\left(\mathbf{x}^{(0)} - \alpha\mathbf{z}\right)$$

at $\alpha_1 = 0$, $\alpha_2 = 0.5$, and $\alpha_3 = 1$ as follows:

$$\boxed{\begin{aligned}
&\alpha_1 = 0, \quad g_1 = 111.975,\\
&\alpha_2 = 0.5, \quad g_2 = 2.53557, \quad h_1 = \frac{g_2 - g_1}{\alpha_2 - \alpha_1} = -218.878,\\
&\alpha_3 = 1, \quad g_3 = 93.5649, \quad h_2 = \frac{g_3 - g_2}{\alpha_3 - \alpha_2} = 182.059, \quad h_3 = \frac{h_2 - h_1}{\alpha_3 - \alpha_1} = 400.937.
\end{aligned}}$$

Thus

$$\boxed{P_2(x) = f[x_0] + f[x_0, x_1](x - x_0) + f[x_0, x_1, x_2](x - x_0)(x - x_1)}$$

$$P(\alpha) = 111.975 - 218.878\alpha + 400.937\alpha(\alpha - 0.5).$$

We have $P'(\alpha) = 0$ when $\alpha = \alpha_0 = 0.522959$. Since $g_0 = g\left(\mathbf{x}^{(0)} - \alpha_0\mathbf{z}\right) = 2.32762$ is smaller than g_1 and g_3, we set

$$\boxed{\mathbf{x}^{(1)}} = \mathbf{x}^{(0)} - \alpha_0\mathbf{z} = \mathbf{x}^{(0)} - 0.522959\mathbf{z} = (0.0112182, 0.0100964, -0.522741)^t$$

and

$$g(\mathbf{x}^{(1)}) = 2.32762.$$

Table 10.5 contains the remainder of the results. A true solution to the nonlinear system is $(0.5, 0, -0.5235988)^t$, so $\mathbf{x}^{(2)}$ would likely be adequate as an initial approximation for Newton's method or Broyden's method. One of these quicker converging techniques would be appropriate at this stage, since 70 iterations of the Steepest Descent method are required to find $\|\mathbf{x}^{(k)} - \mathbf{x}\|_\infty < 0.01$. ■

Table 10.5

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$g(x_1^{(k)}, x_2^{(k)}, x_3^{(k)})$
2	0.137860	−0.205453	−0.522059	1.27406
3	0.266959	0.00551102	−0.558494	1.06813
4	0.272734	−0.00811751	−0.522006	0.468309
5	0.308689	−0.0204026	−0.533112	0.381087
6	0.314308	−0.0147046	−0.520923	0.318837
7	0.324267	−0.00852549	−0.528431	0.287024

Newton's method ① 先畫圖猜 $\vec{P}_0$

Thm 10.7

let $\vec{p} = \vec{G}(\vec{p})$. $\delta > 0$

if ① (i) $\dfrac{\partial g_i(\vec{x})}{\partial x_k}$: contin on $N_\delta = \{\vec{x} \mid \|\vec{x} - \vec{p}\| < \delta\}$

(ii) $\dfrac{\partial g_i}{\partial x_k}(\vec{p}) = 0$

② $\dfrac{\partial^2 g_i}{\partial x_j \partial x_k}(\vec{x})$: contin , with $\left|\dfrac{\partial^2 g_i}{\partial x_j \partial x_k}(\vec{x})\right| \le M$ $\forall \vec{x} \in N_\delta$

$\Rightarrow \exists \hat{\delta} < \delta$, $\forall \vec{P}_0$ s.t. $\|\vec{P}_0 - \vec{P}\| < \hat{\delta}$

$\vec{P}_n = \vec{G}(\vec{P}_{n-1})$ conv. to $\vec{P}$

at least quadratically

Moreover, n : large.

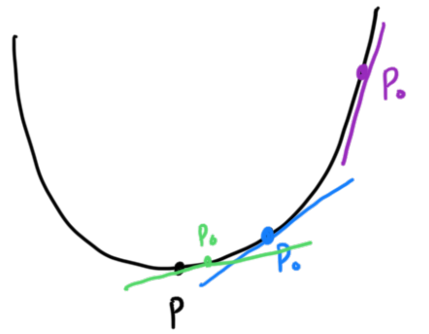
$$\|\vec{P}^n - \vec{P}\|_\infty \le \frac{n^2 M}{2} \|\vec{P}_{n-1} - \vec{P}\|_\infty^2$$

Steepest Descent

converge linearly

$\vec{P}_0$: 隨便 (like Bisection)

$\Delta \|\vec{P}_0 - \vec{P}\|$ 越大 , conv. 越快



Comb.

① ramdon $\vec{P}_0$, steepest descent.

② $g(\vec{P}_k) < \hat{\delta}$, $\hat{\delta}$ from Thm 10.7

③ pick $\vec{X}_0 = \vec{P}_k$, newton method