Lagrange Polynomial in Interpolation

Liyuan Cao

OptML Feb 5, 2020

Interpolation

We have function f and sample points

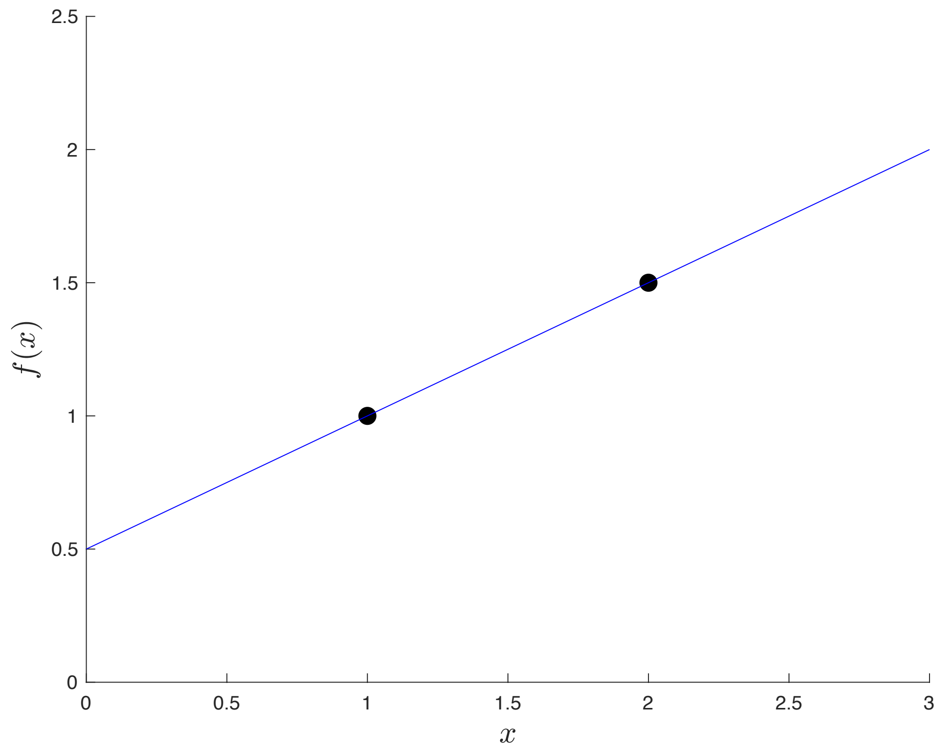
$$Y = \begin{bmatrix}1\\2\end{bmatrix}, f(Y) = \begin{bmatrix}1\\1.5\end{bmatrix}.$$

We want to approximate f with a linear function

$$m(x) = \alpha_0 + \alpha_1 x.$$

We need to solve

$$\begin{bmatrix}1 & 1\\1 & 2\end{bmatrix}\begin{bmatrix}\alpha_0\\\alpha_1\end{bmatrix} = \begin{bmatrix}1\\1.5\end{bmatrix}.$$



Interpolation

We have function f and sample points

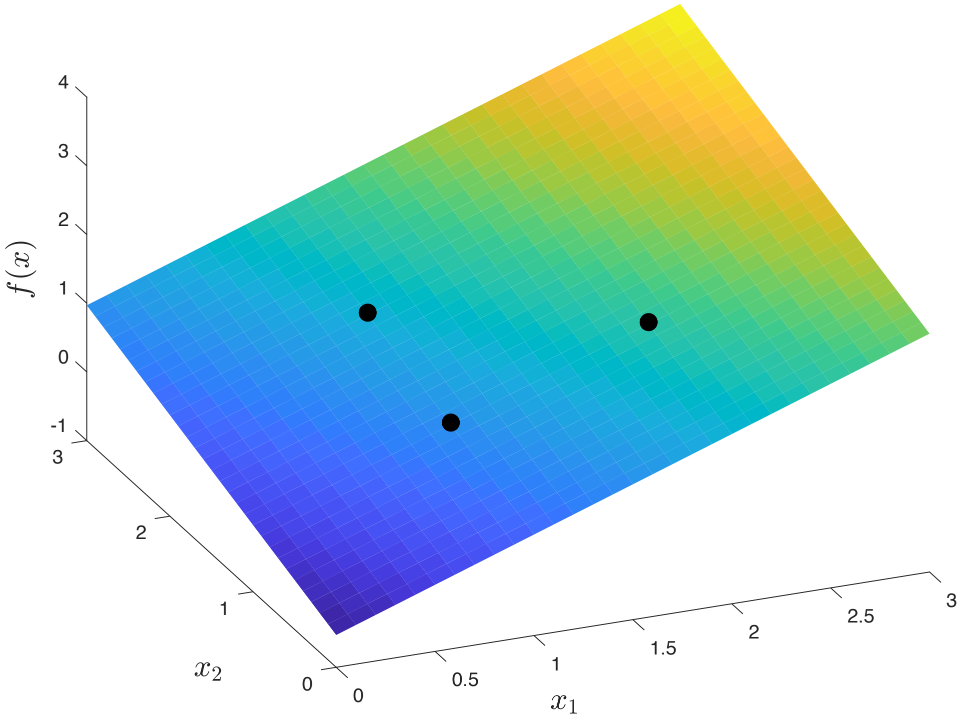
$$Y = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}, f(Y) = \begin{bmatrix} 1 \\ 1.5 \\ 2 \end{bmatrix}.$$

We want to approximate f with a linear function

$$m(x) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2.$$

We need to solve

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1.5 \\ 2 \end{bmatrix}.$$



Interpolation

We have function f and sample points

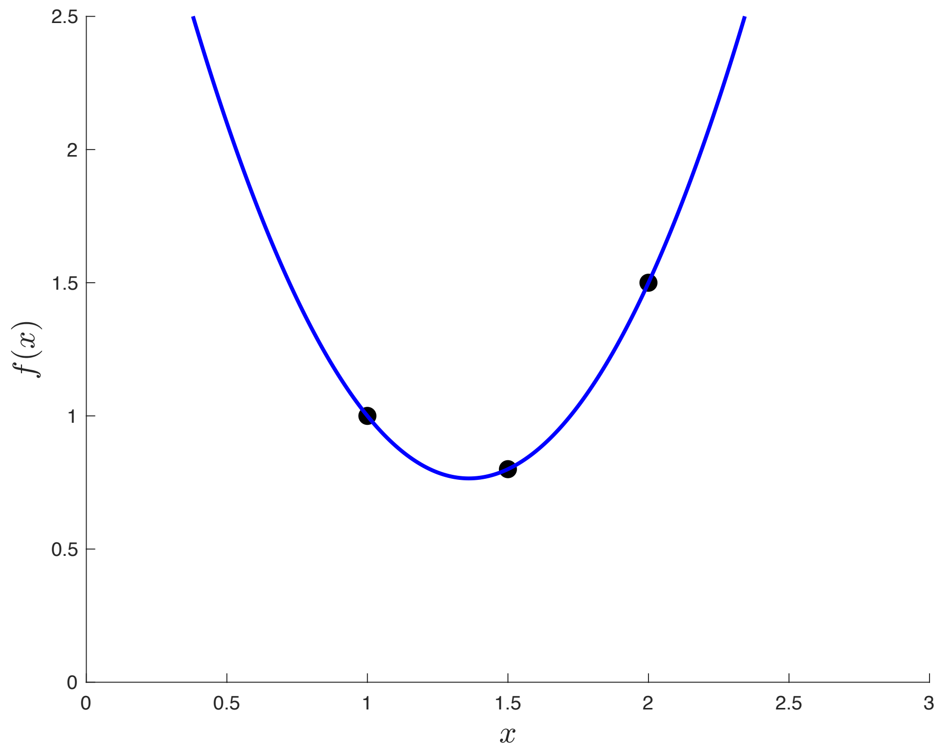
$$Y = \begin{bmatrix} 1 \\ 1.5 \\ 2 \end{bmatrix}, f(Y) = \begin{bmatrix} 1 \\ 0.8 \\ 1.5 \end{bmatrix}.$$

We want to approximate f with a linear function

$$m(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2.$$

We need to solve

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1.5 & 2.25 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.8 \\ 1.5 \end{bmatrix}.$$



DFO Algorithm Using Interpolation

Start with some x^* and some Δ.

LOOP:

1. Get a sample set $\{y_0, y_1, y_2, \cdots, y_p\}$;
2. Calculate the interpolation model $m(x)$;
3. Solve $x_1 \leftarrow \min_{x \in TR} m(x)$;
4. If $f(x_1) \leq f(x^*)$, then $x^* \leftarrow x_1$;
5. Adjust Δ accordingly.

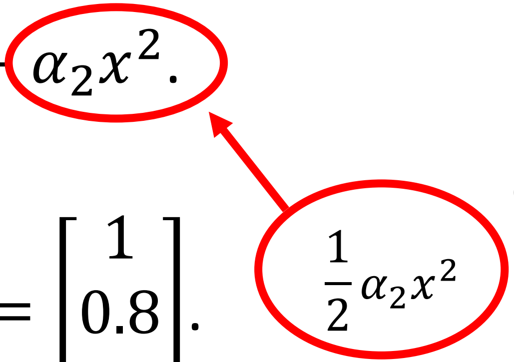
Interpolation

We have function f and sample points

$$Y = \begin{bmatrix} 1 \\ 1.5 \\ 2 \end{bmatrix}, f(Y) = \begin{bmatrix} 1 \\ 0.8 \\ 1.5 \end{bmatrix}.$$

We want to approximate f with a linear function

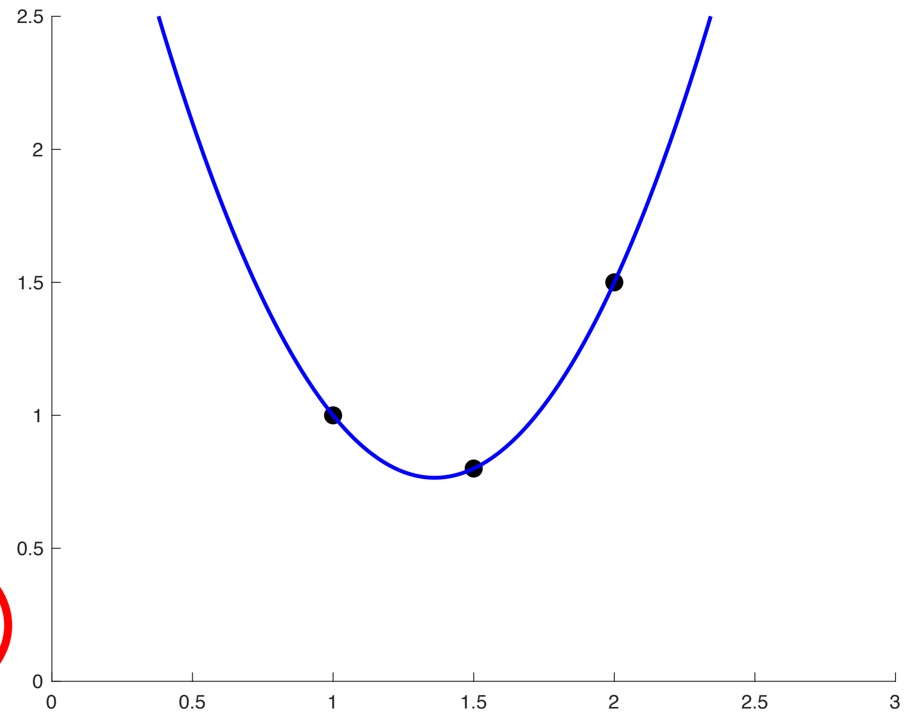
$$m(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2.$$



$$\frac{1}{2}\alpha_2 x^2$$

We need to solve

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1.5 & 2.25 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.8 \\ 1.5 \end{bmatrix}.$$



Basis

Is there a "correct" one?

$$\phi(x_1, x_2) = [1 \quad x_1 \quad x_2 \quad x_1^2 \quad x_1 x_2 \quad x_2^2]^T$$

$$\phi(x_1, x_2) = \left[1 \quad x_1 \quad x_2 \quad \frac{1}{2}x_1^2 \quad x_1 x_2 \quad \frac{1}{2}x_2^2\right]^T$$

NO! And this is also correct.

$$\phi(x_1, x_2) = \left[1 \quad x_1 + x_2 \quad x_2 - x_1 + x_1 x_2 \quad \frac{1}{2}x_1^2 - x_1 \quad x_1 x_2 \quad \frac{1}{2}x_2^2 + 3.14\right]^T$$

Basis

If we call this natural basis

$$\bar{\phi}(x) = \begin{bmatrix} 1 & x_1 & x_2 & \frac{1}{2}x_1^2 & x_1 x_2 & \frac{1}{2}x_2^2 \end{bmatrix}^T,$$

then any basis

$$\phi(x) = P\,\bar{\phi}(x)$$

with non-singular P is a valid basis.

Let $p+1$ be the length of $\bar{\phi}(x)$, then $P \in \mathbb{R}^{(p+1)\times(p+1)}$.

Interpolation

Given the form of $m(x)$, if there are $p + 1$ coefficients, we need $p + 1$ sample points to uniquely define the interpolation model. Let them be

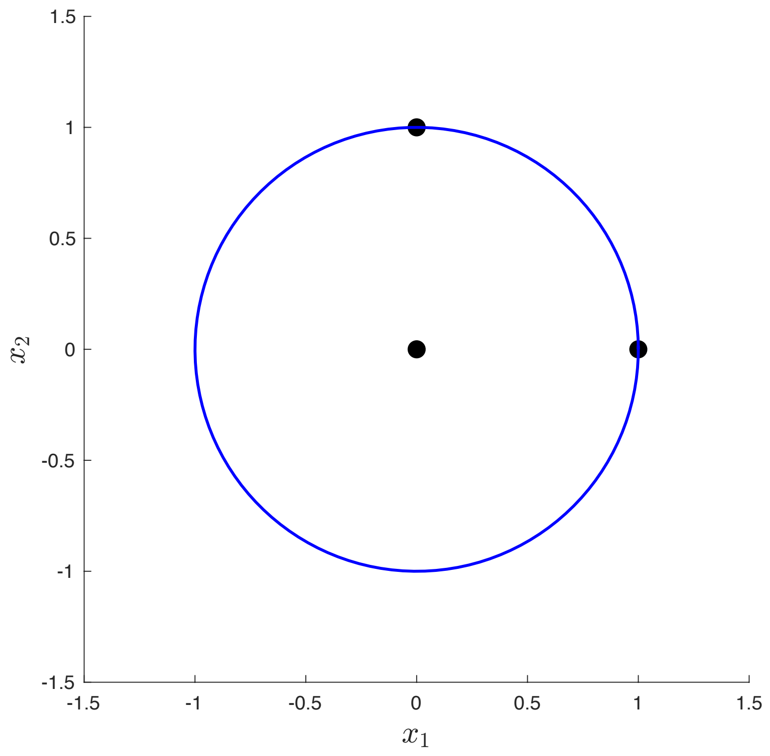
$$\{y_0, y_1, y_2, \cdots, y_p\}.$$

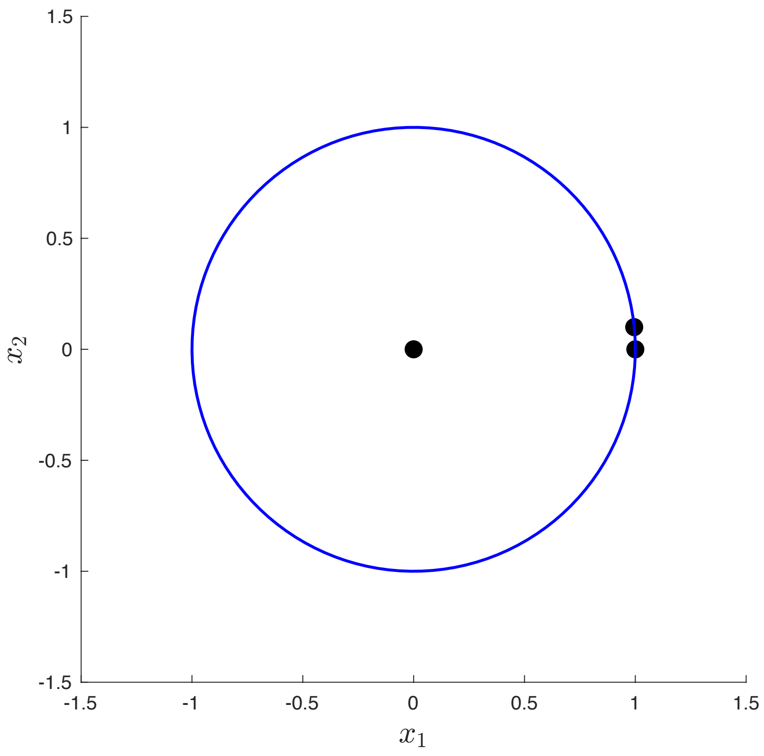
Let $\Phi = \begin{bmatrix} \phi(y_0)^T \\ \phi(y_1)^T \\ \vdots \\ \phi(y_p)^T \end{bmatrix}$, $\vec{f} = \begin{bmatrix} f(y_0) \\ f(y_1) \\ \vdots \\ f(y_p) \end{bmatrix}$, and α be all the coefficients,

then we can get the interpolation model $m(x) = \alpha^T \phi(x)$ by solving

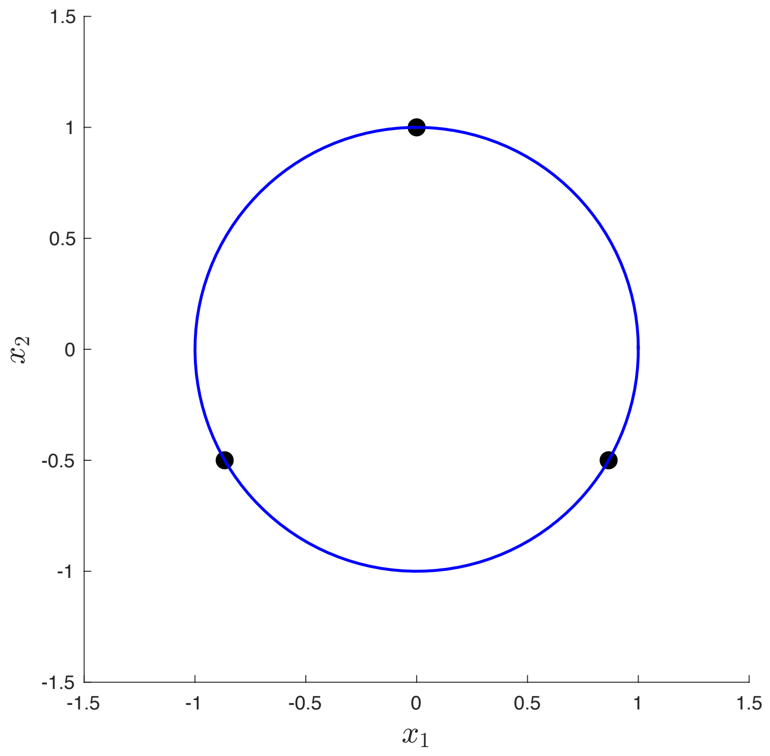
$$\Phi\alpha = \vec{f} \, .$$

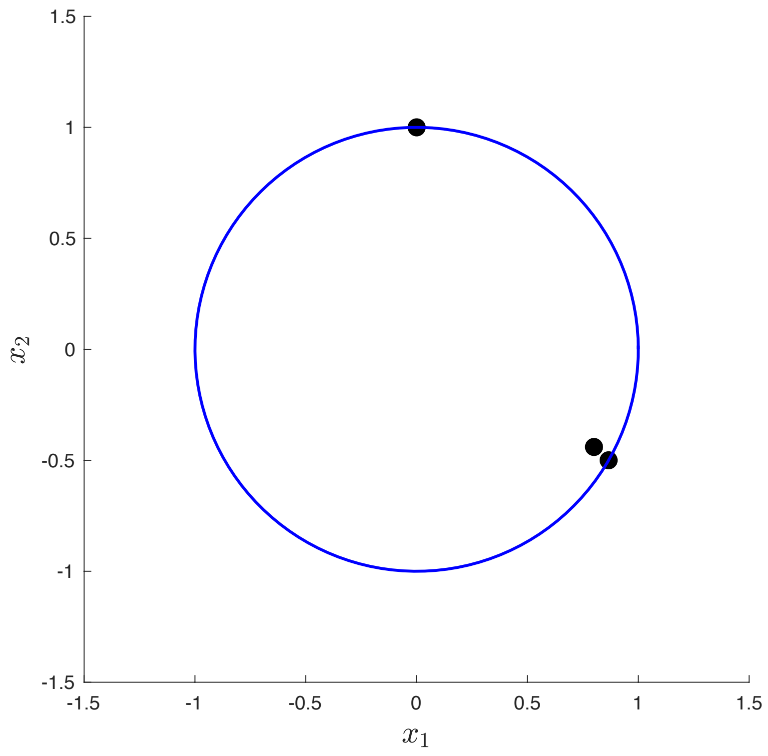
Something to Consider



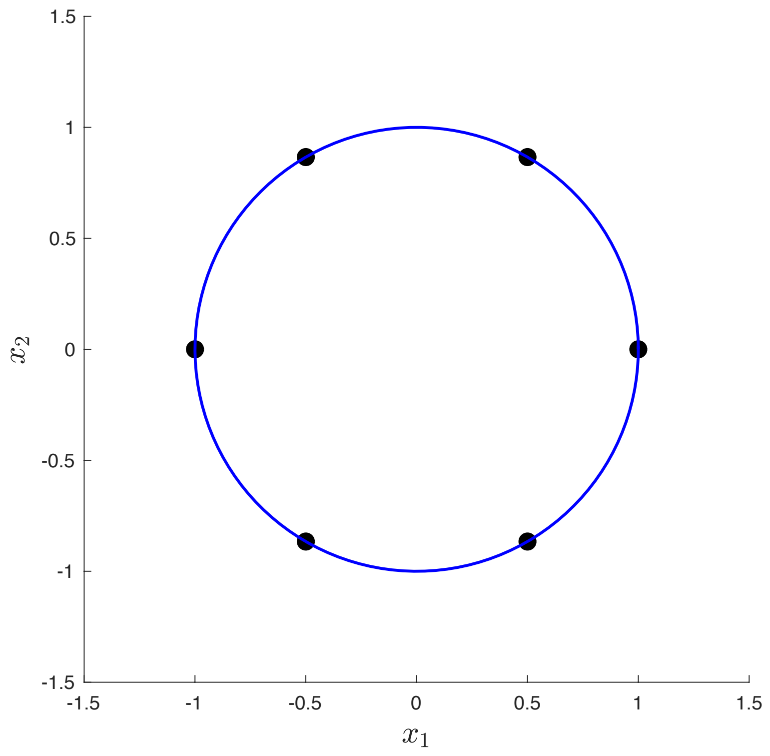


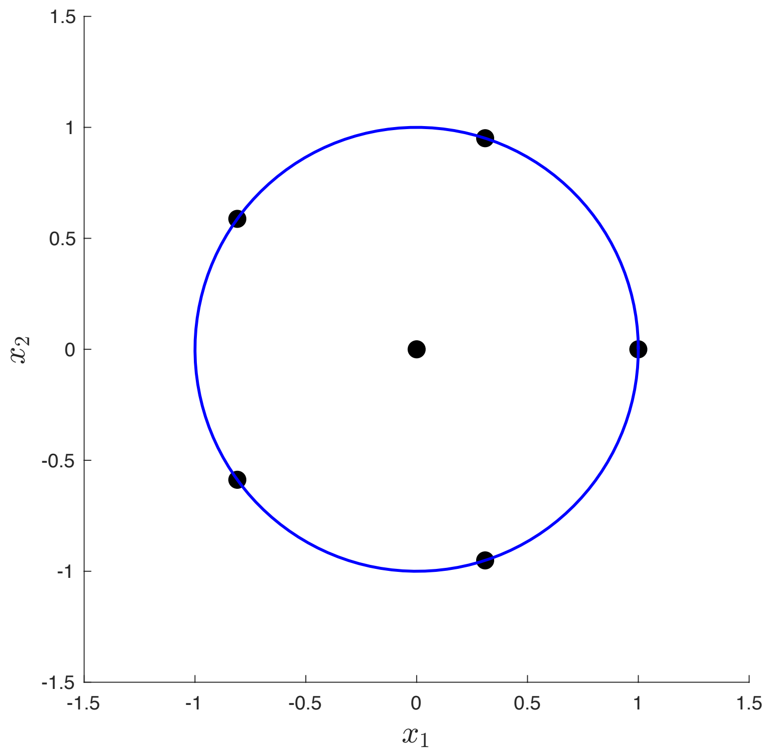
Something to Consider



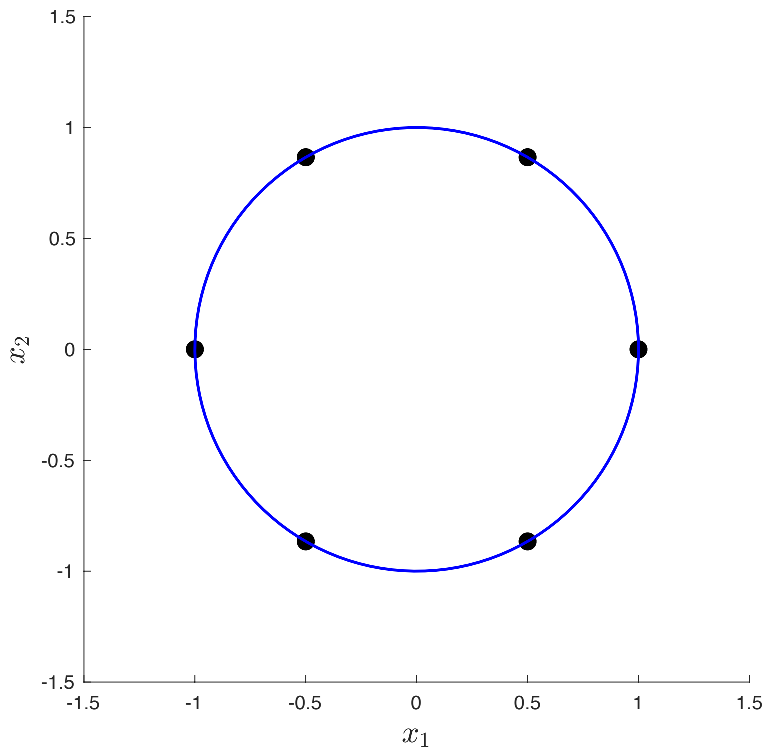


Something to Consider





Something to Consider

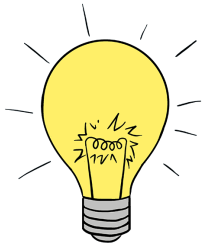


Poisedness

means how well a sample set is distributed in an area for interpolation purpose

- How to measure poisedness exactly (mathematically)?
- How do we improve poisedness?

Poisedness

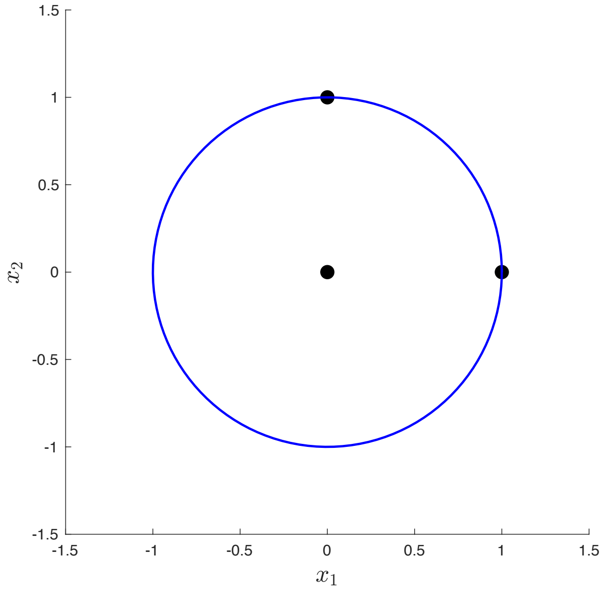


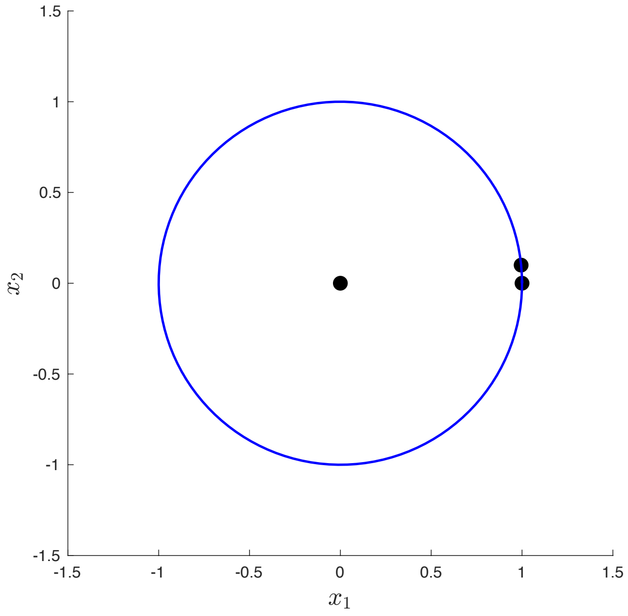
Remember $\Phi\alpha = \vec{f}$?

Let $\text{cond}(\Phi)$ be our measure of poisedness!

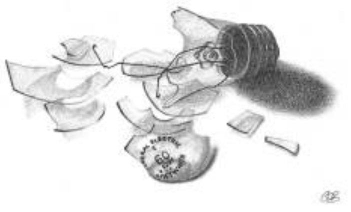
$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \text{cond}(\Phi) = 3.7321$$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0.9950 & 0.0998 \end{bmatrix}, \text{cond}(\Phi) = 30.2128$$





Poisedness

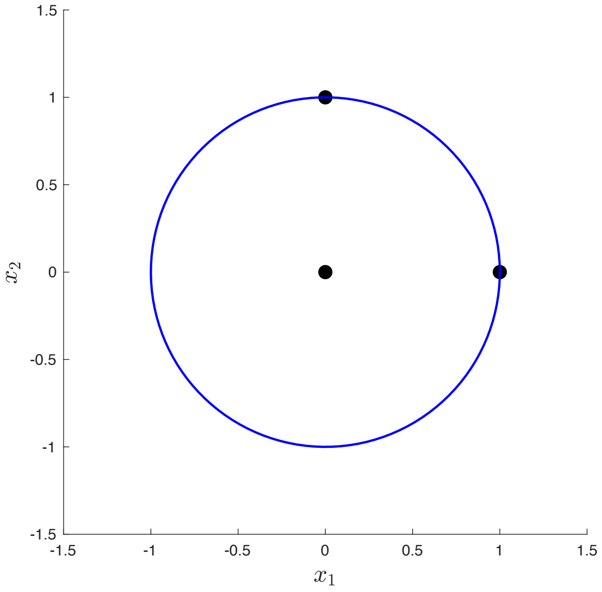


But that really depends on the basis. The condition number can be anything in $[1, \infty)$. Let

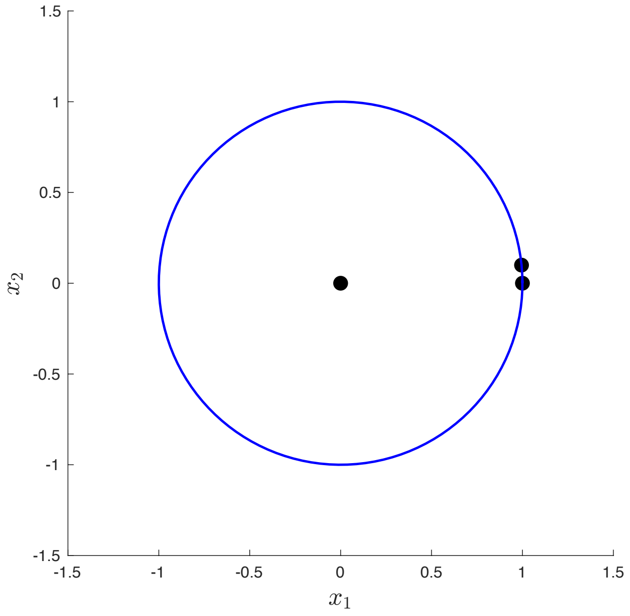
$$\phi(x) = [1 - x_1 - 0.0501x_2, \quad x_1 - 9.9699x_2, \quad 10.0200x_2]^T$$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.9499 & -9.9699 & 10.0200 \end{bmatrix},$$

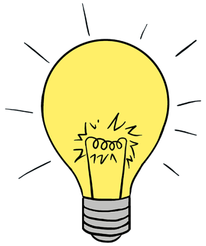
$$\mathrm{cond}(\Phi) = 20.0802$$



$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathrm{cond}(\Phi) = 1$$



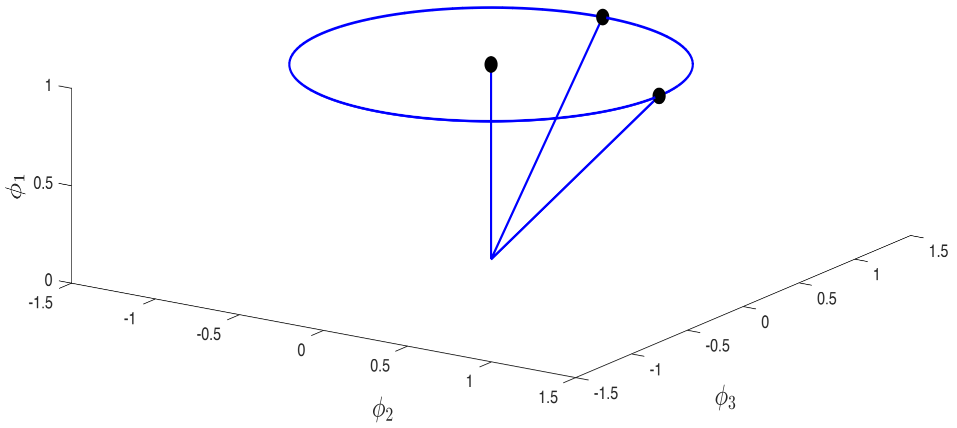
Poisedness

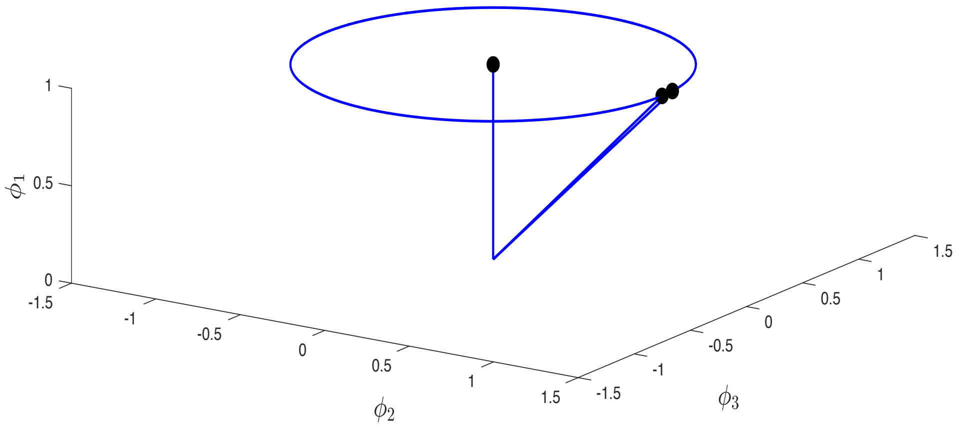


What about $|\Phi|$?

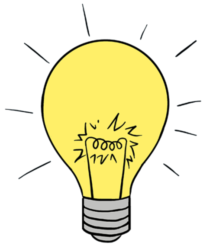
$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, |\Phi| = 1$$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0.9950 & 0.0998 \end{bmatrix}, |\Phi| = 0.0998$$





Poisedness



What about $|\Phi|$? Let

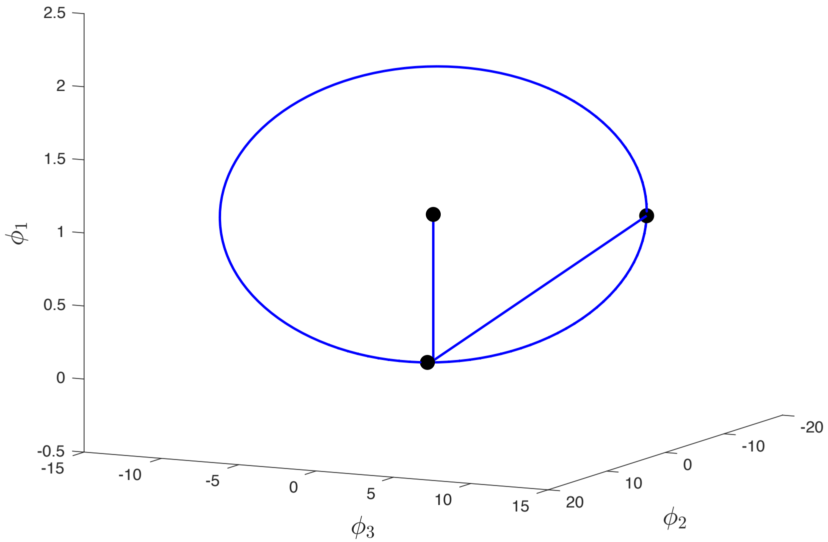
$$\phi(x) = [1 - x_1 - 0.0501x_2, \quad x_1 - 9.9699x_2, \quad 10.0200x_2]$$

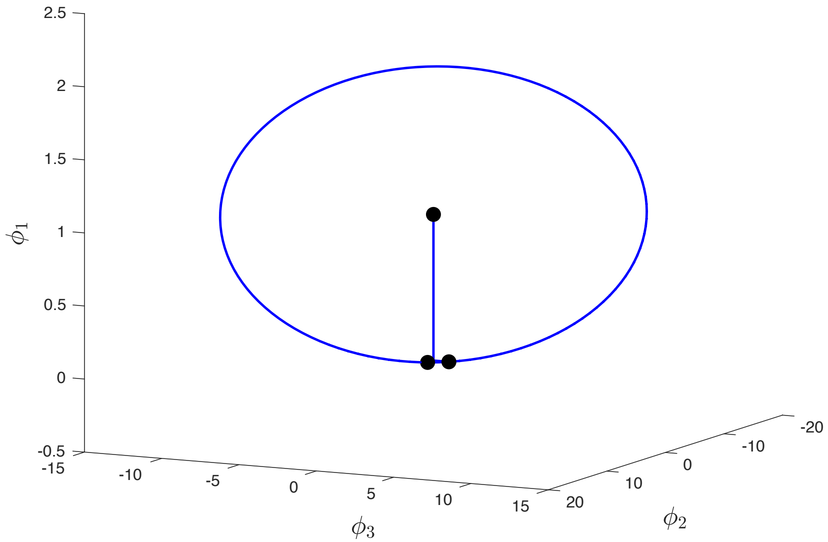
determinant of product = product of determinant

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.9499 & -9.9699 & 10.0200 \end{bmatrix},$$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, |\Phi| = 1$$

$$|\Phi| = 10.0200$$



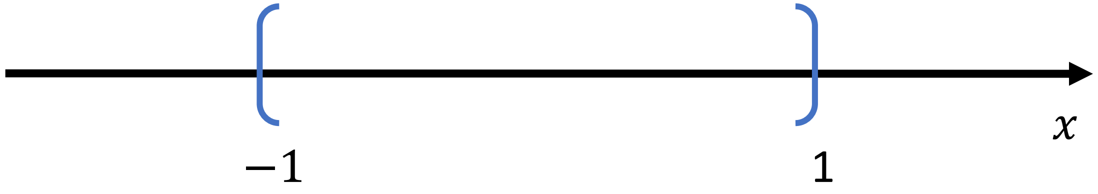


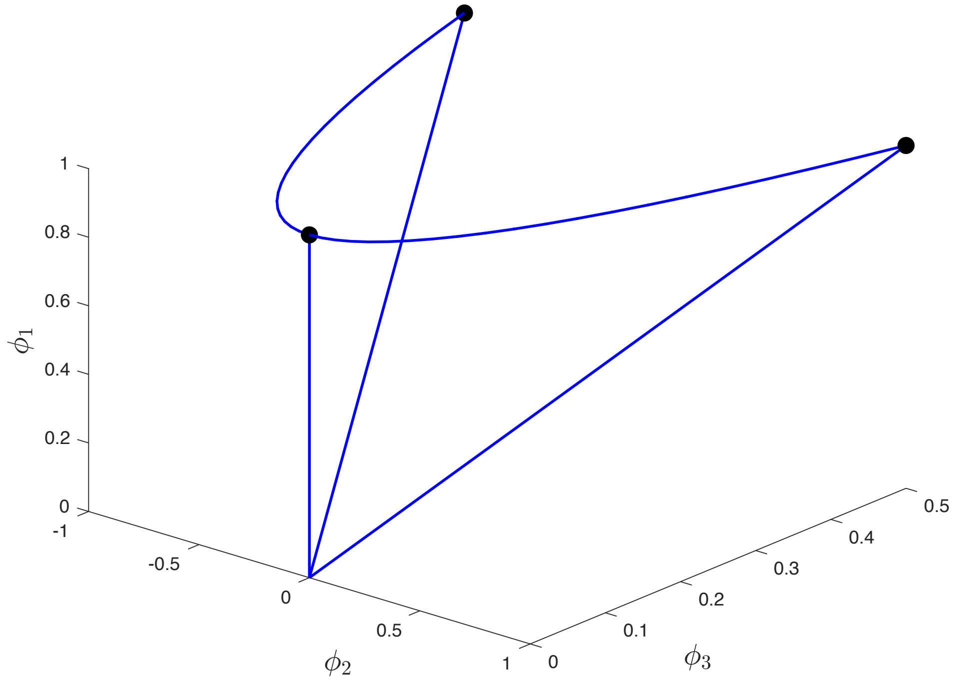
Poisedness

$$\Phi = \begin{bmatrix} 1 & -1 & 0.5 \\ 1 & 0 & 0 \\ 1 & 1 & 0.5 \end{bmatrix}$$

quadratic interpolation on 1D function

$$m(x) = \alpha_0 + \alpha_1 x + \frac{1}{2}\alpha_2 x^2$$





Poisedness

Cool, we just need to find the sample set with max $|\Phi|$,

but this intuition correct?

And what about Lagrange polynomial?

Lagrange Polynomial

Definition 3.3. *Given a set of interpolation points $Y = \{y^0, y^1, \ldots, y^p\}$, a basis of $p_1 = p+1$ polynomials $\ell_j(x)$, $j = 0, \ldots, p$, in $\mathcal{P}_n^d$ is called a basis of Lagrange polynomials if*

$$\ell_j(y^i) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

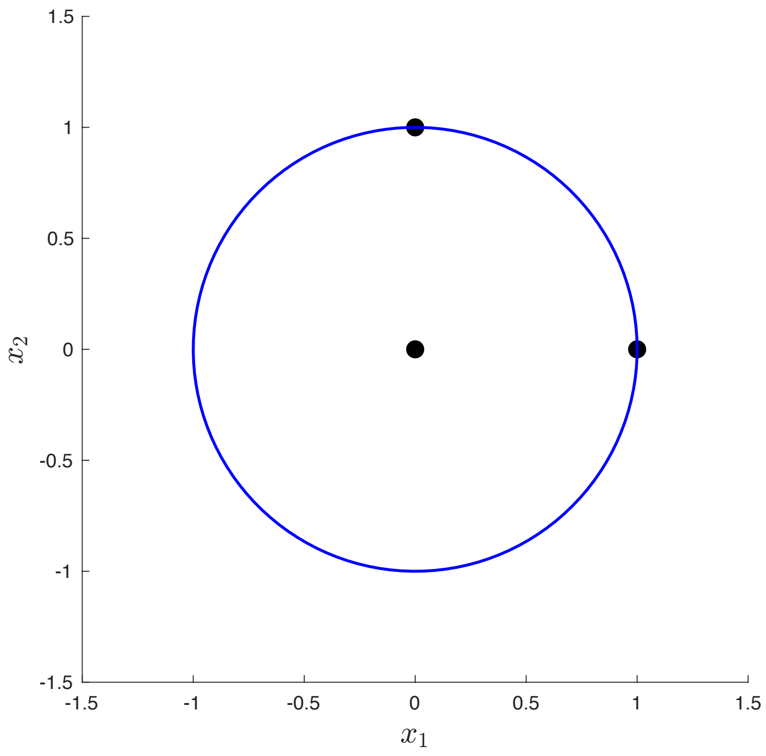
Lemma 3.4. *If Y is poised, then the basis of Lagrange polynomials exists and is uniquely defined.*

Lemma 3.5. *For any function $f : \mathbb{R}^n \to \mathbb{R}$ and any poised set $Y = \{y^0, y^1, \ldots, y^p\} \subset \mathbb{R}^n$, the unique polynomial $m(x)$ that interpolates $f(x)$ on Y can be expressed as*

$$m(x) = \sum_{i=0}^{p} f(y^i)\ell_i(x),$$

where $\{\ell_i(x), i = 0, \ldots, p\}$ is the basis of Lagrange polynomials for Y.

Lagrange Polynomial



sample: $\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \leftarrow \phi(y_0)^T \\ \leftarrow \phi(y_1)^T \\ \leftarrow \phi(y_2)^T \end{matrix}$$

$$A^T = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \leftarrow l_0(x) = 1 - x_1 - x_2 \\ \leftarrow l_1(x) = 0 + x_1 + 0 \\ \leftarrow l_2(x) = 0 + 0 + x_2 \end{matrix}$$

Lagrange Polynomial

Definition 3.3. *Given a set of interpolation points $Y = \{y^0, y^1, \ldots, y^p\}$, a basis of $p_1 = p+1$ polynomials $\ell_j(x)$, $j = 0, \ldots, p$, in $\mathcal{P}_n^d$ is called a basis of Lagrange polynomials if*

$$\ell_j(y^i) = \delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

sample: $\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \leftarrow \phi(y_0)^T \\ \leftarrow \phi(y_1)^T \\ \leftarrow \phi(y_2)^T \end{matrix}$$

$$A^T = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \leftarrow l_0(x) = 1 - x_1 - x_2 \\ \leftarrow l_1(x) = 0 + x_1 + 0 \\ \leftarrow l_2(x) = 0 + 0 + x_2 \end{matrix}$$

Definition 3.3: $\Phi A = I$

Lagrange Polynomial

Lemma 3.4. *If Y is poised, then the basis of Lagrange polynomials exists and is uniquely defined.*

sample: $\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \leftarrow \phi(y_0)^T \\ \leftarrow \phi(y_1)^T \\ \leftarrow \phi(y_2)^T \end{matrix}$$

$$A^T = \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{matrix} \leftarrow l_0(x) = 1 - x_1 - x_2 \\ \leftarrow l_1(x) = 0 + x_1 + 0 \\ \leftarrow l_2(x) = 0 + 0 + x_2 \end{matrix}$$

Lemma 3.4:
If Φ is nonsingular, A is unique.

Lagrange Polynomial

Lemma 3.5. *For any function $f : \mathbb{R}^n \to \mathbb{R}$ and any poised set $Y = \{y^0, y^1, \ldots, y^p\} \subset \mathbb{R}^n$, the unique polynomial $m(x)$ that interpolates $f(x)$ on Y can be expressed as*

$$m(x) = \sum_{i=0}^{p} f(y^i)\ell_i(x),$$

where $\{\ell_i(x), i = 0, \ldots, p\}$ is the basis of Lagrange polynomials for Y.

sample: $\left\{\begin{bmatrix}0\\0\end{bmatrix} \quad \begin{bmatrix}1\\0\end{bmatrix} \quad \begin{bmatrix}0\\1\end{bmatrix}\right\}$

$$\Phi = \begin{bmatrix}1 & 0 & 0\\1 & 1 & 0\\1 & 0 & 1\end{bmatrix} \qquad \begin{matrix}\leftarrow \phi(y_0)^T\\ \leftarrow \phi(y_1)^T\\ \leftarrow \phi(y_2)^T\end{matrix}$$

$$A^T = \begin{bmatrix}1 & -1 & -1\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} \qquad \begin{matrix}\leftarrow l_0(x) = 1 - x_1 - x_2\\ \leftarrow l_1(x) = 0 + x_1 + 0\\ \leftarrow l_2(x) = 0 + 0 + x_2\end{matrix}$$

Lemma 3.5:

$m(x) = \phi(x)^T A\vec{f}$, because

$\Phi\alpha = \vec{f}, \alpha = \Phi^{-1}\vec{f} = A\vec{f}, m(x) = \phi(x)^T\alpha$

Lagrange Polynomial

$$\Phi A = I$$
$$A\Phi = I$$
$$\phi(x)^T A\Phi = \phi(x)^T$$
$$l(x)^T \Phi = \phi(x)^T$$

$$\sum_{i=0}^{p} l_i(x)\phi(y_i) = \phi(x)$$

The $|l_i(x)|$ measures how well the sample set spans $\{\phi(x) | x \text{ in the trust region}\}$

Lagrange Polynomial

$$l(x)^T \Phi = \phi(x)^T$$
$$\Phi^T l(x) = \phi(x)$$

By Cramer's Rule: $l_i(x) = \frac{\left|\Phi^{\mathrm{T}}_{(\mathrm{i})}\right|}{\left|\Phi^T\right|}$, which is

$$\frac{\text{volume when } i\text{th point is replaced with } x}{\text{volume of the original sample set}}.$$

$$\Lambda = \max_{0 \leq i \leq p} \max_{x \in TR} |l_i(x)|$$

Λ is a measure of poisedness, and the volume can be increased by a factor of Λ if we replace y_i with x.

The Theory

Write f in the form of its Taylor expansion about x:

$$f(y_i) = f(x) + \nabla f(x)(y_i - x) + \frac{1}{2!}\nabla^2 f(x)(y_i - x)^2 + \frac{1}{3!}\nabla^3 f(x)(y_i - x)^3 + \cdots,$$

where $\nabla^3 f(x)(y_i - x)^3$ means the inner product of $\nabla^3 f(x)$ and three $(y_i - x)$ vectors.

Let t be the dth order tylor expansion of f:

$$t(y_i) = f(x) + \nabla f(x)(y_i - x) + \cdots + \frac{1}{d!}\nabla^d f(x)(y_i - x)^d.$$

The Taylor series with explicit remainder gives

$$|(f - t)(y_i)| = \frac{1}{(d+1)!}\nabla^{d+1} f(\xi)(y_i - x)^{d+1}$$

for some $\xi = \beta(y_i - x)$ with some $\beta \in [0, 1]$. If $\nabla^d f$ is L_d Lipschitz continuous, then $\|\nabla^{d+1} f(\xi)\| \leq L_d$ and

$$|(f - t)(y_i)| \leq \frac{L_d}{(d+1)!}\|y_i - x\|^{d+1}.$$

$$|(f-t)(y_i)| \leq \frac{L_d}{(d+1)!}\|y_i - x\|^{d+1}.$$

Then for any $0 \leq r \leq d$

$$\begin{aligned}
&\|\nabla^r m(x) - \nabla^r f(x)\| \\
=& \|\nabla^r (m-t)(x) - \nabla^r (f-t)(x)\| \\
=& \|\nabla^r (m-t)(x)\| \\
=& \left\| \sum_{i=0}^{p} (f-t)(y_i) \nabla^r l_i(x) \right\| \\
\leq& \sum_{i=0}^{p} |(f-t)(y_i)| \cdot \|\nabla^r l_i(x)\| \\
\leq& \sum_{i=0}^{p} \frac{L_d}{(d+1)!} \|\nabla^r l_i(x)\| \|y_i - x\|^{d+1}.
\end{aligned}$$

Let $r = 0, d = 2$ and we have

$$\begin{aligned}
&|m(x) - f(x)| \\
\leq& \sum_{i=0}^{p} \frac{L_2}{6} |l_i(x)| \cdot \|y_i - x\|^3 \\
\leq& p \frac{L_2}{6} \Lambda (2\Delta)^3
\end{aligned}$$

where Δ is the radius of the trust region.