Kergin-Lagrange Interpolation

JWR

Sometime Ago

1. For each $k = 0, 1, 2, \dots$ denote by Δ^k the standard k-simplex, the set of all

$$u = (u_1, u_2, \dots, u_k) \in \mathbf{R}^k$$

such that $0 \le u_1 \le u_2 \le \cdots \le u_k \le 1$. A **singular k-simplex** in a topological space V is a continuous map $\sigma : \Delta^k \to V$. It determines a linear functional $\iota(\sigma) : C^0(V) \to \mathbf{R}$ on the space of continuous functions on V via the formula

$$\langle \iota(\sigma), f \rangle = \int_{\Delta^k} f(\sigma(u))\, du$$

for $f \in C^0(V)$; the integral on the right is with respect to the standard measure on $\mathbf{R}^k$. For a constant function the integral is independent of σ:

$$\langle \iota(\sigma), 1 \rangle = \frac{1}{k!}.$$

For a continuous map $\phi : V \to W$ and a function $g \in C^0(W)$ we have the formula

$$\langle \iota(\phi_* \sigma), g \rangle = \langle \iota(\sigma), \phi^* g \rangle .$$

(This formula is a triviality. It is not the change of variables formula for integrals.)

2. From now on V (and eventually W) will denote a finite dimensional vector space over the real numbers $\mathbf{R}$. An **affine singular k simplex** $\sigma : \Delta^k \to V$ has the form

$$\sigma(u) = z_0 + \sum_{j=1}^{k} u_j (z_j - z_0)$$

for $u \in \Delta^k$. The points

$$z_0 = \sigma(0), \quad z_j = \sigma(e_j),$$

where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ (1 in the jth position) are called the **vertices** of σ. (The terminology is somewhat misleading; if the vertices are not in general position, some of them may fail to be extreme points of the image $\sigma(\Delta^k)$.)

3. Denote by $S^k(V)$ the homogeneous polynomials of degree k on V. Via polarization we have the identification

$$S^k(V) = L_s^k(V, \mathbf{R})$$

with the symmetric k-multilinear forms. We denote the inhomogeneous polynomials of degree $\leq r$ by

$$P^r(V) = \bigoplus_{k=0}^{r} S^k(V)$$

The dimensions of $S^k(V)$ and $P^r(V)$ are given by

$$\dim S^k(V) = \binom{k+n-1}{k}, \qquad \dim P^r(V) = \binom{r+n}{r}, \qquad n = \dim V.$$

(The former formula is by the Ehrenfest trick and the latter by the identification $P^r(\mathbf{R}^n) = S^r(\mathbf{R}^{n+1})$.) We take the binomial coefficient $\binom{m}{k}$ to be zero if $m < k$ and $k > 0$ so that $S^k(V) = \{0\}$ if $V = \{0\}$ and $k > 0$. Always however we take $S^0(V) = \mathbf{R}$.

4. Let $Q_k(V)$ denote the vector space of constant coefficient linear differential operators on V which are homogeneous of degree k. The formula

$$q(D)e^{\langle \xi, \cdot \rangle} = q(\xi)e^{\langle \xi, \cdot \rangle}$$

establishes a natural isomorphism

$$Q_k(V) = S^k(V^*).$$

This isomorphism may be described as follows. Let $C^k(V)$ denotes the space of k times continuously differentiable functions on V. Each $q \in S^k(V^*) = S^k(V)^*$ corresponds to the element $q(D) \in Q_k(V)$ defined by

$$(q(D)f)(x) = \left\langle q, D^k f(x) \right\rangle$$

where $x \in V$ and the k-the derivative $D^k f(x) \in L_s^k(V, \mathbf{R}) = S^k(V)$. Our conventions require $Q_0(V) = \mathbf{R}$ for any V and $Q_k(V) = \{0\}$ when $k > 0$ and $V = \{0\}$.

5. A vector $v \in V$ determines $D_v \in Q_1(V)$ via

$$(D_v f)(x) = Df(x)v = \frac{d}{dt}\bigg|_{t=0} f(x + tv)$$

for $x \in V$ and $f \in C^1(V)$. The vector space $Q_k(V)$ is spanned by the k-fold products $D_{v_1} D_{v_2} \cdots D_{v_k}$ as $v_1, v_2, \ldots v_k$ range over V. A function $f \in C^1(V)$ and a vector $v \in V$ determine a vectorfield $fv : V \to V$ satisfying the formula

$$\operatorname{div}(fv) = D_v f$$

for the divergence.

6. Let $\phi : V \to W$ an affine map and $\phi_\# : V \to W$ be its linear part, i.e. $\phi_\#$ is linear and $\phi(x) = \phi(x_0) + \phi_\#(x - x_0)$ for $x, x_0 \in V$. There is an induced transformation

$$\phi_\# : Q_k(V) \to Q_k(W)$$

characterized by

$$\phi_\# D_v = D_{\phi_\# v}$$

for $v \in V$ and

$$\phi_\#\big(q_1(D)q_2(D)\big) = \phi_\#\big(q_1(D)\big)\phi_\#\big(q_2(D)\big)$$

for $q_1(D) \in Q_{k_1}(V)$, $q_2(D) \in Q_{k_2}(V)$. The formula

$$\phi^*\big(\phi_\# q(D))g\big) = q(D)\phi^* g$$

holds for $q(D) \in Q_k(V)$ and $g \in C^k(W)$.

7. Given a singular k-simplex $\sigma : \Delta^k \to V$ and a differential operator $q(D) \in Q_j(D)$ we define a linear functional $\iota(\sigma, q) : C^j(V) :\to \mathbf{R}$ via

$$\langle \iota(\sigma, q), f \rangle = \langle \iota(\sigma), q(D)f \rangle$$

for $f \in C^j(V)$. (In the sequel we only consider those functionals $\iota(\sigma, q)$ for which $k = j$ and where σ is affine.) The formula

$$\langle \iota(\sigma, q), \phi^* g \rangle = \langle \iota(\phi_* \sigma, \phi_* q), g \rangle$$

holds for an affine map $\phi : V \to W$ and function $g \in C^j(W)$.

8. Fix a sequence

$$X = (x_0, x_1, x_2, \dots, x_r) \in V^{r+1}$$

in a vector space V of dimension n. We allow repetitions in X. For $k = 0, 1, 2, \dots, r$ let $X^{(k)}$ denote the set of all affine singular k-simplices σ with vertices $x_{i_0}, x_{i_1}, \dots x_{i_k}$ with $0 \le i_0 < i_1 < \cdots < i_k \le r$; the set $\sigma(\Delta^k) \subset V$ is the convex hull of $x_{i_0}, x_{i_1}, \dots x_{i_k}$.

9. For each $\sigma \in X^{(k)}$ define a subspace $B(\sigma) \subset C^r(V)^*$ by

$$B(\sigma) = \operatorname{span}\{\iota(\sigma, q) : q \in Q_k(V)\}$$

and for $j = -1, 0, 1, 2, \dots, r$ define

$$B_j(X) = \operatorname{span}\{\iota : \iota \in B(\sigma),\ \sigma \in X^{(k)},\ k \le j\}.$$

Define $B(X) = B_r(X)$. There is a filtration

$$\{0\} = B_{-1}(X) \subset B_0(X) \subset B_1(X) \subset \cdots \subset B_r(X) = B(X)$$

and $B(\sigma) \subset B_j(X)$ for $\sigma \in X^{(k)}$ and $k \le j$.

Theorem 10. *There is a direct sum decomposition*

$$C^r(V) = P^r(V) \oplus B(X)^\perp$$

where $B(X)^\perp$ denotes the annihilator of $B(X) \subset C^r(V)^*$.

Proof. It suffices to prove

(A) If $p \in P^r(V)$ is such that $\langle \eta, p \rangle = 0$ for all $\eta \in B(X)$, then $p = 0$, and

(B) $\dim B(X) \le \dim P^r(V)$.

11. *We prove (A).* Let $p \in P^r(V)$ satisfy the hypothesis of (A) and write p in multiindex notation

$$p(y) = \sum_{|\alpha| \le r} p_\alpha y^\alpha$$

with respect to affine coordinates $y_1, y_2, \dots, y_n$ on V. Suppose inductively that $p_\alpha = 0$ for $|\alpha| > k$; we show that $p_\beta = 0$ for $|\beta| = k$. Fix β and let $q(D) = (\partial/\partial y)^\beta$. Then

$$\big(q(D)p\big)(y) = \beta! p_\beta$$

(a constant) by the induction hypothesis. Choose any element $\sigma \in X^{(k)}$. Then

$$0 = \langle \iota(\sigma, q), p \rangle = \frac{\beta! p_\beta}{k!}$$

ao $p_\beta = 0$ as required.

Lemma 12. *Let $\sigma \in X^{(k)}$ and $v \in V$ be parallel to $\sigma(\Delta^k)$, i.e. $v \in \sigma_\#(\mathbf{R}^k)$. Then for any $q(D) \in Q_{k-1}(V)$ we have*

$$\iota(\sigma, vq) \in B_{k-1}(X)$$

where $(vq)(D) \in Q_k(V)$ is the composition

$$(vq)(D) = D_v \circ q(D).$$

Proof. As v is parallel to σ there exists $w \in \mathbf{R}^k$ with $\sigma_\# w = v$. Let $g = q(D)f$ and id : $\Delta^k \to \Delta^k$ denote the identity map. Then

$$\begin{aligned}
\langle \iota(\sigma, vq), f \rangle &= \langle \iota(\sigma), D_v g \rangle \\
&= \langle \iota(\mathrm{id}), \sigma^*(D_v g) \rangle \\
&= \langle \iota(\mathrm{id}), (D_w(\sigma^* g)) \rangle \\
&= \langle \iota(\mathrm{id}), \mathrm{div}(\sigma^* g) w) \rangle \\
&= \sum_\tau (w \cdot \hat{\tau}) \, \langle \iota(\mathrm{id}), \tau * g \rangle \\
&= \sum_\tau (w \cdot \hat{\tau}) \, \langle \iota(\tau, q), f \rangle
\end{aligned}$$

where the penultimate step is by the divergence theorem, τ ranges over the faces of σ, and $\hat{\tau}$ denotes the outward normal to τ. We have shown that

$$\iota(\sigma, vq) = \sum_\tau (w \cdot \hat{\tau} \iota(\tau, q)$$

which proves the lemma.

13. Identify V and V^* via an inner product. For $\sigma \in X^{(k)}$ let $\sigma^\perp \subset V$ denote the vector subspace perpendicular to the simplex $\sigma(\Delta^k)$, i.e. the vectors in $\sigma^\perp$ and the vectors in $\sigma_\#(\mathbf{R}^k)$ are orthogonal. The inclusion $\sigma^\perp \subset V$ induces an inclusion $Q_k(\sigma^\perp) \subset Q_k(V)$ The lemma gives a direct sum decomposition

$$B(\sigma) = B(\sigma, \perp) \oplus B(\sigma) \cap B_{k-1}(X) \qquad (\diamondsuit)$$

where

$$B(\sigma, \perp) = \{\iota(\sigma, q) : q \in Q_k(\sigma^\perp)\}.$$

To check this choose a basis of $Q_k(V)$ consisting consisting of compositions $D_{v_1} D_{v_2} \cdots D_{v_k}$ where each v_i is either parallel or perpendicular to $\sigma(\Delta^k)$. If any v_i is parallel to $\sigma(\Delta^k)$, the corresponding functional $\iota(\sigma, q)$ lies in $B(\sigma) \cap B_{k-1}(X)$ by the lemma. Those compositions where all v_i are perpendicular to $\sigma(\Delta^k)$ lie in $B(\sigma, \perp)$ by definition.

14. *We prove (B).* We may assume w.l.o.g. that

$$\dim \sigma(\Delta^k) = k \qquad \text{for } \sigma \in X^{(k)} \text{ and } k \leq n \tag{$\heartsuit$}$$

where $n = \dim V$. This is because the set of all $X \in V^{r+1}$ for which ($\heartsuit$) holds is dense (and open) in V^{r+1} and $\dim B(X)$ is a lower semicontinuous function of X.

By ($\heartsuit$) we have

$$\dim \sigma^\perp = n - k$$

for $\sigma \in X^{(k)}$ and $k \leq n$. Hence

$$\dim B(\sigma, \perp) \leq \dim Q_k(\sigma^\perp) = \binom{k + (n-k) - 1}{k} = \binom{n-1}{k}.$$

The set $X^{(k)}$ has cardinality $\binom{r+1}{k+1} = \binom{r+1}{r-k}$ so

$$\dim B_k(X)/B_{k-1}(X) \leq \binom{r+1}{r-k}\binom{n-1}{k}.$$

As the subspaces $B_k(X)$ filter $B(X)$ we may sum these inequalities to obtain

$$\dim B(X) \leq \sum_{k=0}^{r} \binom{r+1}{r-k}\binom{n-1}{k} = \binom{r+n}{r} = \dim P^r(V)$$

as required. (In the last step we used the equation

$$\sum_{k=0}^{r} \binom{a}{r-k}\binom{b}{k} = \binom{a+b}{r}$$

with $a = r + 1$ and $b = n - 1$. This formula says that the hypergeometric probabilities sum to one.)

Remark 15. It follows that $\dim B(X) = \dim P^r(V)$. The subspaces $B(\sigma)$ span $B(X)$ by definition so by ($\diamondsuit$) and induction the spaces $B(\sigma, \perp)$ span $B(X)$. Under the nondegeneracy hypothesis ($\heartsuit$) the proof gives a direct sum decomposition

$$B(X) = \bigoplus_{k=0}^{n-1} \bigoplus_{\sigma \in X^{(k)}} B(\sigma, \perp).$$

(If the sum were not direct, the dimension on the left would be smaller than the dimension on the right.)

Definition 16. The projection

$$I_X : C^r(V) \to P^r(V)$$

corresponding to the splitting in theorem 10 is called **Kergin-Lagrange interpolation**. For $f \in C^r(V)$ the polynomial $I_X f$ is the unique polynomial satisfying

$$\langle \iota(\sigma, q), f \rangle = \langle \iota(\sigma, q), I_X f \rangle$$

for every $k = 0, 1, 2, \ldots, r$ every $\sigma \in X^{(k)}$, and every $q(D) \in Q_k(V)$.

Remark 17. By the previous remark, under the nondegeneracy hypothesis ($\heartsuit$), the polynomial $I_X f$ is determined by the derivatives of f of order $< n$, but in general, higher derivatives are required. For example, in the extreme case $x_0 = x_1 = \cdots x_r$, $I_X f$ is the Taylor polynomial of f at the point x_0. In case $n = 1$ the nondegeneracy hypothesis says that all the points x_i are distinct, so that $I_X f$ is the unique polynomial p of degree r such that $p(x_i) = f(x_i)$ for $i = 0, 1, \ldots, r$, i.e. I_X is the Lagrange interpolant of f.

Proposition 18. *Suppose that $\phi : V \to W$ is affine. Then*

$$\phi^* I_{\phi(X)} g = I_X \phi^* g$$

for $g \in C^r(W)$.

Proof. As both sides of the equation are elements of $P^r(V)$ it suffices to show they give the same value at each element $\iota(\sigma, q)$ of $B(X)$. The calculation is

$$\begin{aligned}
\left\langle \iota(\sigma, q), \phi^* I_{\phi(X)} g \right\rangle &= \left\langle \iota(\phi_* \sigma, \phi_* q), I_{\phi(X)} g \right\rangle \\
&= \langle \iota(\phi_* \sigma, \phi_* q), g \rangle \\
&= \langle \iota(\sigma, q), \phi^* g \rangle \\
&= \langle \iota(\sigma, q), I_X \phi^* g \rangle
\end{aligned}$$

as required.

Proposition 19. *Assume that $f \in C^r(V)$ and $q(D) \in Q_k(V)$ with $k \leq r$. Then*

$$q(D)f = 0 \implies q(D)I_X f = 0.$$

Proof. Assume that $q(D)f = 0$ and write

$$I_X f = p_0 + p_1 + p_2 + \cdots + p_r$$

with $p_j \in S^j(V)$. Assume inductively that $q(D)p_i = 0$ for $i > j$; we will show that $q(D)p_j = 0$. For $j < k$ this is automtatic, so assume that $j \geq k$. Using affine coordinates and mutiindex notation write

$$q(D)p_j(y) = \sum_{|\beta|=j-k} b_\beta y^\beta.$$

For $|\alpha| = j - k$ the induction hypothesis gives

$$q_\alpha(D)p = \alpha! b_\alpha$$

where $q_\alpha(D) \in Q_j(V)$ is defined by

$$q_\alpha(D) = D^\alpha q(D), \qquad D^\alpha = \frac{\partial^{|\alpha|}}{\partial y^\alpha}.$$

We have $q_\alpha(D)f = 0$ as $q(D)f = 0$. Choose $\sigma \in X^{(j)}$. Then

$$0 = \langle \iota(\sigma, q_\alpha), f\rangle = \langle \iota(\sigma, q_\alpha), p\rangle = \frac{b_\alpha \alpha!}{(j-k)!}$$

so $b_\alpha = 0$ so (as this holds for all α) $q(D)p_j = 0$ as required.

Corollary 20. *Let L be a vector subspace of V and suppose that $f \in C^r(V)$. Then if f is constant on the translates of L the same is true of $I_X f$.*

Proof. The hypothesis is that $D_v f = 0$ for $v \in L$ and the conclusion is that $D_v I_X f = 0$ for $v \in L$. (This proof will be reused in the proof of theorem 25 below.) We can also use proposition 18: a function f is constant on the translates of L iff $f = \phi^* g$ for some $g \in W = V/L$ where $\phi : V \to W$ is the projection.

Remark 21. The corollary says that if $y_1, y_2, \ldots, y_n$ are affine coordinates and f depends only on the first m of these coordinates, then the same is true of $I_X f$.

Theorem 22. *The map I_X is the unique linear map $I : C^r(V) \to P^r(X)$ such that (i) I is continuous (in the topology of uniform converegence of derivatives of order $\leq r$ on compact sets) and (ii) for any linear functional $\xi : V \to \mathbf{R}$ and any $g \in C^r(\mathbf{R})$ the polynomial $I\xi^* g$ is the Lagrange polynomial interpolating g at the points $\xi(x_0), \xi(x_1), \ldots, \xi(x_r)$.*

Proof. The map I_X satisfies (ii) by proposition 18:

$$I_X \xi^* g = \xi^* I_{\xi X} g.$$

The map I_X satisfies (i) since the functionals $\iota(\sigma, q)$ are continuous. For uniqueness assume that I satisfies (i) and (ii). By (i) and the fact that the polynomials are dense it suffices that $If = I_X f$ for any polynomial. By proposition 18 and remark 17 condition (ii) says that $If = I_X f$ for any function $f \in C^r(V)$ of form $f = \xi^* g$ where $g \in C^r(\mathbf{R})$. Hence by linearity it suffices to show that (for any m) the polynomials of form $f = \xi^* g$ where $\xi \in V^*$ and $g \in P^m(\mathbf{R})$ span $P^m(V)$. If $g(t) = \sum_{k=0}^m g_k t^k$, then $\xi^* g = \sum_{k=0}^m g_k p_{k\xi}$ where

$$p_{k\xi}(x) = \langle \xi, x \rangle^k.$$

Hence it suffices to prove the following

Lemma 23. *The vector space $P^m(V)$ is spanned by the polynomials $p_{k\xi}$ ($k = 0, 1, 2, \ldots, m$, $\xi \in V^*$).*

Proof. By the multinomial formula

$$p_{k\xi}(x) = \sum_{|\alpha|=k} \binom{k}{\alpha} \xi^\alpha x^\alpha.$$

Suppose that $\ell \in P^m(V)^*$ annihilates all $p_{k\xi}$. Let

$$\langle \ell, p \rangle = \sum_{|\alpha| \leq m} \ell_\alpha p_\alpha, \qquad p(x) = \sum_{|\alpha| \leq m} p_\alpha x^\alpha.$$

Then

$$0 = \langle \ell, p_{k\xi} \rangle = \sum_{|\alpha|=k} \ell_\alpha \binom{k}{\alpha} \xi^\alpha$$

for all ξ (and k) so that $\ell_\alpha = 0$ for all α so $\ell = 0$ as required.

Definition 24. A map $I : C^r(V) \to C^r(V)$ satisfies the **GMVP** (**G**eneralized **M**ean **V**alue **P**roperty) iff for every $k = 0, 1, \dots, r$, every $q(D) \in Q_k(V)$, and every choice $0 \le i_0 < i_1 < i_2 < \cdots i_k \le r$ of distimct indices there is a point $\bar{x}$ in the convex hull of the points $x_{i_0}, x_{i_1}, x_{i_2}, \dots, x_{i_k}$ such that

$$q(D)f(\bar{x}) = q(D)If(\bar{x}).$$

Theorem 25. *The map I_X is the unique map $I : C^r(V) \to P^r(X)$ which (i) is linear and (ii) satisfies the GMVP for X.*

Proof. To see that I_X satisfies the GMVP for X note that the convex hull of $x_{i_0}, x_{i_1}, \dots, x_{i_k}$ is $\sigma(\Delta^k)$ for the corresponding $\sigma \in X^{(k)}$. Thus (ii) says that *for every for every $k = 0, 1, 2, \dots, r$, every $q(D) \in Q_k(V)$, and for every $\sigma \in X^{(k)}$ the functions $q(D)f$ and $(D)I_X f$ agree at some point $\bar{x}$ of $\sigma(\Delta^k)$.* The equation

$$\langle \iota(\sigma, q), f \rangle = \langle \iota(\sigma, q), I_X f \rangle$$

takes the form

$$\int_{\Delta^k} (q(D)f)(\sigma(u))\, du = \int_{\Delta^k} (q(D)I_X f)(\sigma(u))\, du.$$

Now if g_1 and g_2 are real valued continuous functions on a connected set which have the same integral over that set then there must be a point in that set where they are equal: otherwise, one would be greater than the other at every point and the integrals would not be equal. Thus $q(D)f(\bar{x}) = q(D)I_X f(\bar{x})$ at some $\bar{x} = \sigma(\bar{x})$ as required.

To prove uniqueness assume that I satisfies (i) and (ii) of theorem 25; we prove that I satisfies (i) and (ii) of theorem 22.

Step 1. *Theorem 25 is true whem $n = 1$.* The GMVP says that f and If agree to order $m_k - 1$ where m_k is the number of i such that $x_i = x_k$ (so that $m_k = 1$ when the elements of X are distinct). The Lagrange interpolant of f is the unique polynomial of degree $\le r$ with this propoerty.

Step 2. *Proposition 19 (and hence also corollary 20) remains true when I is read for I_X.* The proof is essentially the same: Assume that $q(D)f = 0$ and write

$$If = p_0 + p_1 + p_2 + \cdots + p_r$$

with $p_j \in S^j(V)$. Assume inductively that $q(D)p_i = 0$ for $i > j$; we will show that $q(D)p_j = 0$. For $j < k$ this is automtatic, so assume that $j \ge k$. Using mutiindex notation write

$$q(D)p_j(y) = \sum_{|\beta|=j-k} b_\beta y^\beta.$$

For $|\alpha| = j - k$ the induction hypothesis gives

$$q_\alpha(D)p = \alpha! b_\alpha$$

where $q_\alpha(D) = D^\alpha q(D)$ as the the proof of proposition 19. We have $q_\alpha(D)f = 0$ as $q(D)f = 0$. Choose $\sigma \in X^{(j)}$. By the GMVP we have

$$0 = q_\alpha(D)f(\bar{x}) = q_\alpha(D)If(\bar{x}) = \alpha! b_\alpha$$

for some $\bar{x} = \sigma(\bar{u})$ in the $\sigma(\Delta^k)$ so $b_\alpha = 0$ so (as this holds for all α) $q(D)p_j = 0$ as required.

Step 3. *The map I satisfies condition (ii) of theorem 22.* In other words, for any $\xi \in V^*$ we have $I\xi^* g = \xi^* I_{\xi(X)} g$ for $g \in C^r(\mathbf{R})$. By step 2 (the analog of corollary 20) each $g \in C^r(V)$ determines a polynomial $p \in P^r(\mathbf{R})$ with $\xi^* p = I\xi^* g$; define $(\xi_* I) : C^r(\mathbf{R}) \to P^r(\mathbf{R})$ by $(\xi_* I)g = p$. Then $\xi^*(\xi_* I) = I\xi^*$ so $\xi_* I$ satisfies the GMVP. Hence by step 1 we have $I\xi^* = \xi^* I_{\xi(X)}$ as required.

Step 4. *The map I is continuous.* We write If in multiindex notation:

$$(If)(x) = \sum_{|\alpha| \le r} (I_\alpha f) x^\alpha;$$

We must shows that each of the linear functionals I_α is continuous. Assume inductively that this is true for $|\alpha| > k$; we show it is true for $|\alpha| = k$. Apply $D^\alpha = \partial^{|\alpha|}/\partial x^\alpha$ to obtain

$$D^\alpha (If)(x) = \alpha! I_\alpha f + (R_\alpha f)(x)$$

where

$$(R_\alpha f)(x) = \sum_{\beta > \alpha} \frac{\beta!}{(\beta - \alpha)!} (I_\beta f) x^{\beta - \alpha}.$$

By the induction hypothesis there is a large compact set (which might as well be the convex hull of X) and a constant C such that

$$|(R_\alpha f)(x)| \le C \|f\|_r$$

where

$$\|f\|_r = \sup\{|D^\gamma f(x) : |\gamma| \le r,\ x \in K\}.$$

Choose $\sigma \in X^{(k)}$. By the GMVP there is an $\bar{x} \in \sigma(\Delta^k)$ such that

$$|D^\alpha(If)(\bar{x})| = |D^\alpha f(\bar{x})| \le \|f\|_r.$$

Hence

$$|I_\alpha f| = \frac{|D^\alpha f(\bar{x}) - R_\alpha f(\bar{x})|}{\alpha!} \le \frac{C+1}{\alpha}\|f\|_r$$

as required.

References

[1] P. Kergin: A natural interpolation of C^k functions, J. Approximation Theory, **19** (1980) 278-293.

[2] S. Helgason: *The Radon Transform*, Birkhäuser, 1980.

[3] C.A. Micchelli: Algebraic aspects of interpolation, in *Approximation Theory*, Proceedings of Symposia in Applied Mathematics, **36** AMS (1986) pp. 81-102.