Computing the Discrete Gauss Transform

Pramook Khungurn

December 11, 2019

Given N Gaussian distributions located at $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ with weights $q_1, q_2, \dots, q_N$. The **discrete Gauss transform** evaluated at point $\mathbf{x}$ is given by

$$G(\mathbf{x}) = \sum_{i=1}^{N} q_i e^{-\|\mathbf{x}-\mathbf{x}_i\|^2/h^2}$$

The positions of the Gaussians are called the **source points**, and $\mathbf{x}$ is called the **target point**. This document is written as I study algorithms to compute the discrete Gauss transforms.

The main note I consult is "The fast Gauss transform with all the proofs" by Vikas C. Raykar [1].

1 Hermite Polynomial and Functions

- The **Hermite polynomial** $H_n(y)$ is defined as:

$$H_n(y) = (-1)^n e^{y^n} \frac{\mathrm{d}^n}{\mathrm{d}y^n} e^{-y^2}.$$

- The generating functiong for the Hermite polynomial is:

$$e^{2yx-x^2} = \sum_{n=0}^{\infty} \frac{x^n}{n!} H_n(y).$$

- Multiplying boths sides by e^{-y^2} yields:

$$e^{-(y-x)^2} = \sum_{n=0}^{\infty} \frac{x^n}{n!} h_n(y)$$

  where $h_n(y)$ is the **Hermite function**. The function is defined as:

$$h_n(y) = e^{-y^2} H_n(y).$$

- Putting in the bandwidth and expanding the function around c, we have

$$e^{-(y-x)^2/h^2} = e^{-[(y-c)/h-(x-c)/h]^2} = \sum_{n=0}^{\infty} \frac{1}{n!} \bigg(\frac{x-c}{h} \bigg)^n h_n \bigg(\frac{y-c}{h} \bigg).$$

- The following recurrence relation is useful in evaluating the Hermite function:

$$h_{n+1}(y) = 2y h_n(y) - 2n h_{n-1}(y).$$

- Hermite polynomial satisfies the Cramer's inequality:

$$|H_n(y)| \leq K 2^{n/2} \sqrt{n!} e^{y^2/2}$$

where $K < 1.09$. This gives the following bound on the Hermite function:

$$\frac{1}{n!}|h_n(y)| \leq K 2^{n/2} \frac{1}{\sqrt{n!}} e^{-y^2/2}.$$

However, the following version is also true:

$$\frac{1}{n!}|h_n(y)| \leq 2^{n/2} \frac{1}{\sqrt{n!}} e^{-y^2/2}.$$

2 Multi-Dimensional Expansion of the Gaussian Kernel

- A **multi-index** $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ is a d-tuple of non-negative integers.

  The length of α, denoted by $|\alpha|$, is deinfed to be $\alpha_1 + \cdots + \alpha_d$.

  We say that $\alpha \geq p$ if $\alpha_i \geq i$ for all i. The proposition $\alpha \leq p$ is defined similarly.

  The factorial of α, denoted by $\alpha!$, is defined to be $\alpha_1! \alpha_2! \cdots \alpha_d!$.

  The d-variate monomial $\mathbf{x}^\alpha$ is defined to be $x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_d^{\alpha_d}$. Notice that the degree of x^α is $|\alpha|$.

  The αth derivative with respect to $\mathbf{x}$ is

$$\frac{\mathrm{d}^\alpha}{\mathrm{d}\mathbf{x}^\alpha} = \frac{\partial^{\alpha_1}}{\partial x^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x^{\alpha_2}} \cdots \frac{\partial^{\alpha_d}}{\partial x^{\alpha_d}}.$$

- The multi-dimentional Hermite function is define as:

$$h_\alpha(\mathbf{y}) = e^{-\|\mathbf{y}\|^2} H_\alpha(y) = h_{\alpha_1}(y_1) h_{\alpha_2}(y_2) \cdots h_{\alpha_d}(y_d).$$

- The Hermite expansion of $e^{-\|\mathbf{x}-\mathbf{y}\|^2}$ is given by:

$$e^{-\|\mathbf{x}-\mathbf{y}\|^2/h^2} = \sum_{\alpha \geq 0} \frac{1}{\alpha!} \left(\frac{\mathbf{x}-\mathbf{c}}{h} \right)^\alpha h_\alpha \left(\frac{\mathbf{y}-\mathbf{c}}{h} \right).$$

  The Taylor expansion is given by:

$$e^{-\|\mathbf{x}-\mathbf{y}\|^2/h^2} = \sum_{\beta \geq 0} \frac{1}{\beta!} h_n \left(\frac{\mathbf{x}-\mathbf{c}}{h} \right) \left(\frac{\mathbf{y}-\mathbf{c}}{h} \right)^\beta.$$

- The Taylor expansion of the multi-dimensional Hermite function around $\mathbf{c}$ is given by:

$$h_\alpha(\mathbf{y}) = \sum_{\beta \geq 0} \frac{(\mathbf{y}-\mathbf{c})^\beta}{\beta!} \frac{\mathrm{d}^\beta}{\mathrm{d}\mathbf{y}^\beta} h_\alpha(\mathbf{c})$$

  where

$$h_\alpha(\mathbf{c}) = (-1)^\alpha \frac{\mathrm{d}^\alpha}{\mathrm{d}\mathbf{y}^\alpha} e^{-\|\mathbf{c}\|^2}.$$

  Because

$$\frac{\mathrm{d}^\beta}{\mathrm{d}\mathbf{y}^\beta} h_\alpha(\mathbf{c}) = \sum_{\beta \geq 0} (-1)^\beta h_{\alpha+\beta}(\mathbf{c}),$$

  we have that

$$h_\alpha(\mathbf{y}) = \sum_{\beta \geq 0} \frac{(\mathbf{y}-\mathbf{c})^\beta}{\beta!} (-1)^\beta h_{\alpha+\beta}(\mathbf{c}).$$

3 Far Field Expansion

- Let B be a box B of side length at most $h/\sqrt{2}$. Let $\mathbf{c}_B$ be the box's center. For any point $\mathbf{y}$, we have that
$$G(\mathbf{y}) = \sum_{i:\mathbf{x}_i \in B} q_i \sum_{\alpha \geq 0} \frac{1}{\alpha!} \left(\frac{\mathbf{x}_i - \mathbf{c}_B}{h} \right)^\alpha h_\alpha \left(\frac{\mathbf{y} - \mathbf{c}_B}{h} \right) \approx \sum_{i:\mathbf{x}_i \in B} q_i \sum_{0 \leq \alpha \leq p} \frac{1}{\alpha!} \left(\frac{\mathbf{x}_i - \mathbf{c}_B}{h} \right)^\alpha h_\alpha \left(\frac{\mathbf{y} - \mathbf{c}_B}{h} \right).$$
Hence, if we define the **moment**
$$A_\alpha^B = \sum_{i:\mathbf{x}_i \in B} q_i \frac{1}{\alpha!} \left(\frac{\mathbf{x}_i - \mathbf{c}_B}{h} \right)^\alpha,$$
then we have
$$G(\mathbf{y}) \approx \sum_{0 \leq \alpha \leq p} A_\alpha^B h_\alpha \left(\frac{\mathbf{y} - \mathbf{c}_B}{h} \right).$$
This is the far-field expansion used in the original fast Gauss transform by Greengard and Strain.

- However, the Greengard expansion is not suitable for high-dimensional Gauss transform because there are $(p+1)^d$ coefficients, which grows exponentially in d.

- Yang et al.[2] proposes another expansion of the Gaussian kernel so that there are $O(p^d)$ terms instead. We shall discuss this expansion here.

- We have that
$$e^{-\|\mathbf{y}-\mathbf{x}_i\|^2/h^2} = e^{-\|(\mathbf{y}-\mathbf{c}_B)-(\mathbf{x}_i-\mathbf{c}_B)\|^2/h^2} = e^{-\|\mathbf{y}-\mathbf{c}_B\|^2/h^2} e^{-\|\mathbf{x}_i-\mathbf{c}_B\|^2/h^2} e^{2(\mathbf{y}-\mathbf{c}_B)\cdot(\mathbf{x}_i-\mathbf{c}_B)/h^2}.$$
We shall show that there are functions Φ_α and Ψ_α such that
$$e^{2(\mathbf{y}-\mathbf{c}_B)\cdot(\mathbf{x}-\mathbf{c}_B)/h^2} = \sum_{\alpha \geq 0} \Phi_\alpha \left(\frac{\mathbf{y} - \mathbf{c}_B}{h} \right) \Psi_\alpha \left(\frac{\mathbf{x}_i - \mathbf{c}_B}{h} \right).$$
Then, we can write the Gaussian kernel as:
$$\begin{aligned} e^{-\|\mathbf{y}-\mathbf{x}_i\|^2/h^2} &= e^{-\|\mathbf{y}-\mathbf{c}_B\|^2/h^2} e^{-\|\mathbf{x}_i-\mathbf{c}_B\|^2/h^2} \sum_{\alpha \geq 0} \Phi_\alpha \left(\frac{\mathbf{y} - \mathbf{c}_B}{h} \right) \Psi_\alpha \left(\frac{\mathbf{x}_i - \mathbf{c}_B}{h} \right) \\ &= \sum_{\alpha \geq 0} \left(e^{-\|\mathbf{y}-\mathbf{c}_B\|^2/h^2} \Phi_\alpha \left(\frac{\mathbf{y} - \mathbf{c}_B}{h} \right) \right) \left(e^{-\|\mathbf{x}_i-\mathbf{c}_B\|^2/h^2} \Psi_\alpha \left(\frac{\mathbf{x}_i - \mathbf{c}_B}{h} \right) \right) \end{aligned}$$
As a result, the Gauss transform can be written as:
$$G(\mathbf{y}) = \sum_{\alpha \geq 0} \left(\sum_{i:\mathbf{x}_i \in B} q_i e^{-\|\mathbf{x}_i-\mathbf{c}_B\|^2/h^2} \Psi_\alpha \left(\frac{\mathbf{x}_i - \mathbf{c}_B}{h} \right) \right) \left(e^{-\|\mathbf{y}-\mathbf{c}_B\|^2/h^2} \Phi_\alpha \left(\frac{\mathbf{y} - \mathbf{c}_B}{h} \right) \right)$$
We can thus define the moment A_α^B as
$$A_\alpha^B = \sum_{i:\mathbf{x}_i \in B} q_i e^{-\|\mathbf{x}_i-\mathbf{c}_B\|^2/h^2} \Psi_\alpha \left(\frac{\mathbf{x}_i - \mathbf{c}_B}{h} \right)$$
and get
$$G(\mathbf{y}) = \sum_{\alpha \geq 0} A_\alpha^B \left(e^{-\|\mathbf{y}-\mathbf{c}_B\|^2/h^2} \Phi_\alpha \left(\frac{\mathbf{y} - \mathbf{c}_B}{h} \right) \right).$$

- We have that

$$(\mathbf{x}\cdot\mathbf{y})^n = \sum_{|\alpha|=n} \frac{n!}{\alpha!}\mathbf{x}^\alpha\mathbf{y}^\alpha.$$

Hence,

$$e^{\mathbf{x}\cdot\mathbf{y}} = \sum_{n=0}^{\infty}\frac{(\mathbf{x}\cdot\mathbf{y})^n}{n!} = \sum_{n=0}^{\infty}\sum_{|\alpha|=n}\frac{1}{\alpha!}\mathbf{x}^\alpha\mathbf{y}^\alpha = \sum_{\alpha\geq 0}\frac{1}{\alpha!}\mathbf{x}^\alpha\mathbf{y}^\alpha.$$

As a result

$$e^{2(\mathbf{x}_i-\mathbf{c}_B)\cdot(\mathbf{y}-\mathbf{c}_B)/h^2} = \sum_{\alpha\geq 0}\frac{2^{|\alpha|}}{\alpha!}\left(\frac{\mathbf{x_i}-\mathbf{c}_B}{h}\right)^\alpha\left(\frac{\mathbf{y}-\mathbf{c}_B}{h}\right)^\alpha.$$

Therefore, we can define

$$\Phi_\alpha(\mathbf{x}) = \frac{2^{|\alpha|}}{\alpha!}\mathbf{x}^\alpha,$$
$$\Psi_\alpha(\mathbf{y}) = \mathbf{y}^\alpha.$$

- This expansion gives:

$$A_\alpha^B = \frac{2^{|\alpha|}}{\alpha!}\sum_{i:\mathbf{x}_i\in B} q_i e^{-\|\mathbf{x}_i-\mathbf{c}_B\|^2/h^2}\left(\frac{\mathbf{x}_i-\mathbf{c}_B}{h}\right)^\alpha,$$

and

$$G(\mathbf{y}) = e^{-\|\mathbf{y}-\mathbf{c}_B\|^2/h^2}\sum_{\alpha\geq 0}A_\alpha^B\left(\frac{\mathbf{y}-\mathbf{c}_B}{h}\right)^\alpha.$$

- Yang et al. truncate the series to terms such that $|a| \leq p$:

$$G(\mathbf{y}) = e^{-\|\mathbf{y}-\mathbf{c}_B\|^2/h^2}\sum_{0\leq\alpha\leq p}A_\alpha^B\left(\frac{\mathbf{y}-\mathbf{c}_B}{h}\right)^\alpha.$$

References

[1] Vikas C. Raykar. The fast gauss transform with all the proofs. `http://www.umiacs.umd.edu/~vikas/publications/FGT.pdf`. Accessed: 08/16/2012.

[2] Changjiang Yang, Ramani Duraiswami, Nail A. Gumerov, and Larry Davis. Improved fast gauss transform and efficient kernel density estimation. In *In ICCV*, pages 464–471, 2003.