DIC: Homework Assignment H4

Team #11: Camilo Martínez 7057573, Honglu Ma 7055053

November 22, 2023

Problem H4.1 (Anisotropic Diffusion Modelling)

(a)

By definition of the diffusion tensor from lecture slides 7, we can state:

$$D(\nabla u_\sigma) := (v_1 \mid v_2)\, diag(g(\mu_1), 1)\, (v_1 \mid v_2)^\top$$

Where v_1 and v_2 are eigenvectors of $J_\rho(\nabla u_\sigma)$ and μ_1 is the larger eigenvalue of $J_\rho(\nabla u_\sigma)$.

(b)

When $\rho \to 0$, we have $J_{\rho\to 0}(\nabla u_\sigma) = K_{\rho\to 0} * (\nabla u_\sigma\, \nabla u_\sigma^\top) = \nabla u_\sigma\, \nabla u_\sigma^\top$. We know from Classroom Work C4.1 that the two eigenvalues of the resulting matrix are $\lambda_1 = 0$ and $\lambda_2 = ||\nabla u_\sigma||^2$; the two eigenvectors are $v_1 = \nabla u_\sigma^\perp$ and $v_2 = \nabla u_\sigma$ respectively. Since $\lambda_2 > \lambda_1$, $\mu_1 = \lambda_1$. With that, we construct our diffusion filter $D(\nabla u_\sigma) = \lambda_1 v_1 v_1^\top + \lambda_2 v_2 v_2^\top = \frac{\nabla u_\sigma^\perp}{||\nabla u_\sigma||} \frac{(\nabla u_\sigma^\perp)^\top}{||\nabla u_\sigma||} + g(||\nabla u_\sigma||^2)\, \frac{\nabla u_\sigma}{||\nabla u_\sigma||}\, \frac{\nabla u_\sigma^\top}{||\nabla u_\sigma||}$.

We can derive the diffusion filter:

$$\begin{aligned}
\partial_t u &= div((\frac{\nabla u_\sigma^\perp}{||\nabla u_\sigma||} \frac{(\nabla u_\sigma^\perp)^\top}{||\nabla u_\sigma||} + g(||\nabla u_\sigma||^2)\, \frac{\nabla u_\sigma}{||\nabla u_\sigma||}\, \frac{\nabla u_\sigma^\top}{||\nabla u_\sigma||})\nabla u_\sigma) \\
&= div(\frac{\nabla u_\sigma^\perp}{||\nabla u_\sigma||} \frac{(\nabla u_\sigma^\perp)^\top}{||\nabla u_\sigma||}\, \nabla u_\sigma + g(||\nabla u_\sigma||^2)\, \frac{\nabla u_\sigma}{||\nabla u_\sigma||}\, \frac{\nabla u_\sigma^\top}{||\nabla u_\sigma||}\, \nabla u_\sigma) \\
&= div(0 + g(||\nabla u_\sigma||^2)\, \frac{\nabla u_\sigma ||\nabla u_\sigma||^2}{||\nabla u_\sigma||^2}) \\
&= div(g(||\nabla u_\sigma||^2)\, \nabla u_\sigma)
\end{aligned}$$

which is the same diffusion filter for non-linear isotropic diffusion.

(c)

When ρ is large and the contrast parameter λ is small, the effect on an image would be an overall smoothing effect. For a large ρ, the kernel K_ρ would be more spread out, therefore taking into account more neighbouring pixels for the convolution. In other words, a large ρ implies a more globalized effect, instead of a more localized one. Furthermore, if λ is small, it implies that the filter is less sensitive to intensity variations, leading to more isotropic diffusion. Since the algorithm becomes less sensitive to edges and high-contrast features, it tends to smooth the image more uniformly instead of taking into account regions with high-contrast variations (edges). Smoothing then becomes more isotropic, affecting both high and low-contrast regions similarly. Both parameters will then lead to the image being smoothed more. In practical terms, this results in an overall blurring effect.

In an image with a pattern of bright and dark stripes, the diffusion process will smooth along the edges of the stripes. Within regions of uniform intensity, the diffusion process will be more pronounced, leading to smoother and less textured areas. We would have a tendency for more uniform smoothing across the entire image and fine image structures, such as the stripes in the pattern, would be blurred out overall.

Problem H4.2 (Directional Splitting of Anisotropic Diffusion)

First, we consider the right hand of the equation:

$$\sum_{i=0}^{3} \partial_{\boldsymbol{e_i}}(w_i \partial_{\boldsymbol{e_i}} u)$$

Where we know that $\partial_{\boldsymbol{n}} u = \boldsymbol{n}^\top \nabla u$, the directional diffusivities w_0, w_1, w_2, w_3 are given by

$$w_0 = a - \delta,\ w_1 = \delta + b,\ w_2 = c - \delta,\ w_3 = \delta - b$$

And the directions are given by

$$\boldsymbol{e_0} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \boldsymbol{e_1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \boldsymbol{e_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \boldsymbol{e_3} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

For $i = 0$, we have:

$$\partial_{\boldsymbol{e_0}}(w_0 \partial_{\boldsymbol{e_0}} u) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^\top \nabla \left[(a - \delta) \begin{pmatrix} 1 \\ 0 \end{pmatrix}^\top \begin{pmatrix} \partial_x u \\ \partial_y u \end{pmatrix} \right] = \partial_x(a \partial_x u) - \partial_x(\delta \partial_x u)$$

Similarly, for $i = 2$, we get:

$$\partial_{\boldsymbol{e_2}}(w_2 \partial_{\boldsymbol{e_2}} u) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}^\top \nabla \left[(c - \delta) \begin{pmatrix} 0 \\ 1 \end{pmatrix}^\top \begin{pmatrix} \partial_x u \\ \partial_y u \end{pmatrix} \right] = \partial_y(c \partial_y u) - \partial_y(\delta \partial_y u)$$

For $i = 1$:

$$\begin{aligned}
\partial_{\boldsymbol{e_1}}(w_1 \partial_{\boldsymbol{e_1}} u) &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}^\top \nabla \left[\tfrac{1}{\sqrt{2}}(\delta + b) \begin{pmatrix} 1 \\ 1 \end{pmatrix}^\top \begin{pmatrix} \partial_x u \\ \partial_y u \end{pmatrix} \right] \\
&= \frac{1}{2}\partial_x(\delta \partial_x u) + \frac{1}{2}\partial_x(b \partial_x u) + \frac{1}{2}\partial_x(\delta \partial_y u) + \frac{1}{2}\partial_x(b \partial_y u) + \frac{1}{2}\partial_y(\delta \partial_x u) + \frac{1}{2}\partial_y(b \partial_x u) + \frac{1}{2}\partial_y(\delta \partial_y u) + \frac{1}{2}\partial_y(b \partial_y u)
\end{aligned}$$

Finally, for $i = 3$:

$$\begin{aligned}
\partial_{\boldsymbol{e_3}}(w_3 \partial_{\boldsymbol{e_3}} u) &= \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}^\top \nabla \left[\tfrac{1}{\sqrt{2}}(\delta - b) \begin{pmatrix} -1 \\ 1 \end{pmatrix}^\top \begin{pmatrix} \partial_x u \\ \partial_y u \end{pmatrix} \right] \\
&= \frac{1}{2}\partial_x(\delta \partial_x u) - \frac{1}{2}\partial_x(b \partial_x u) - \frac{1}{2}\partial_x(\delta \partial_y u) + \frac{1}{2}\partial_x(b \partial_y u) - \frac{1}{2}\partial_y(\delta \partial_x u) + \frac{1}{2}\partial_y(b \partial_x u) + \frac{1}{2}\partial_y(\delta \partial_y u) - \frac{1}{2}\partial_y(b \partial_y u)
\end{aligned}$$

Summing up the terms for $i = 1$ and $i = 3$, we get:

$$\partial_{\boldsymbol{e_2}}(w_2 \partial_{\boldsymbol{e_2}} u) + \partial_{\boldsymbol{e_3}}(w_3 \partial_{\boldsymbol{e_3}} u) = \partial_x(\delta \partial_x u) + \partial_x(b \partial_y u) + \partial_y(b \partial_x u) + \partial_y(\delta \partial_y u)$$

Then, summing up the resulting terms with the ones obtained for $i = 0$ and $i = 2$, we get:

$$\begin{aligned}
\sum_{i=0}^{3} \partial_{\boldsymbol{e_i}}(w_i \partial_{\boldsymbol{e_i}} u) &= \partial_{\boldsymbol{e_0}}(w_0 \partial_{\boldsymbol{e_0}} u) + \partial_{\boldsymbol{e_1}}(w_1 \partial_{\boldsymbol{e_1}} u) + \partial_{\boldsymbol{e_2}}(w_2 \partial_{\boldsymbol{e_2}} u) + \partial_{\boldsymbol{e_3}}(w_3 \partial_{\boldsymbol{e_3}} u) \\
&= \partial_x(\delta \partial_x u) + \partial_x(b \partial_y u) + \partial_y(b \partial_x u) + \partial_y(\delta \partial_y u) \\
&+ \partial_x(a \partial_x u) - \partial_x(\delta \partial_x u) + \partial_y(c \partial_y u) - \partial_y(\delta \partial_y u) \\
&= \partial_x(a \partial_x u) + \partial_x(b \partial_y u) + \partial_y(b \partial_x u) + \partial_y(c \partial_y u)
\end{aligned} \tag{1}$$

On the other hand, let us consider the following derivation which uses the mathematical definition of the divergence of a vector:

$$\begin{aligned}\mathbf{div}\left(\begin{pmatrix}a & b\\ b & c\end{pmatrix}\nabla u\right) &= \mathbf{div}\left(\begin{pmatrix}a & b\\ b & c\end{pmatrix}\begin{pmatrix}\partial_x u\\ \partial_y u\end{pmatrix}\right)\\ &= \mathbf{div}\begin{pmatrix}a\partial_x u + b\partial_y u\\ b\partial_x u + c\partial_y u\end{pmatrix}\\ &= \partial_x(a\partial_x u) + \partial_x(b\partial_y u) + \partial_y(b\partial_x u) + \partial_y(c\partial_y u)\end{aligned} \tag{2}$$

Comparing (1) and (2) term by term, we see that they are equal. Therefore,

$$\sum_{i=0}^{3}\partial_{\boldsymbol{e_i}}(w_i\partial_{\boldsymbol{e_i}}u) = \mathbf{div}\left(\begin{pmatrix}a & b\\ b & c\end{pmatrix}\nabla u\right)$$

Problem H4.3 (δ-Stencil for Isotropic Diffusions)

Problem H4.4 (Anisotropic Diffusion)

Problem H4.5

Give an appropriate positive constant c such that $f(n) \leq c \cdot g(n)$ for all $n > 1$.

1. $f(n) = n^2 + n + 1$, $g(n) = 2n^3$
2. $f(n) = n\sqrt{n} + n^2$, $g(n) = n^2$
3. $f(n) = n^2 - n + 1$, $g(n) = n^2/2$

Solution
We solve each solution algebraically to determine a possible constant c.

Part One

$$\begin{aligned}n^2 + n + 1 &=\\ &\leq n^2 + n^2 + n^2\\ &= 3n^2\\ &\leq c \cdot 2n^3\end{aligned}$$

Thus a valid c could be when $c = 2$.

Part Two

$$\begin{aligned}n^2 + n\sqrt{n} &=\\ &= n^2 + n^{3/2}\\ &\leq n^2 + n^{4/2}\\ &= n^2 + n^2\\ &= 2n^2\\ &\leq c \cdot n^2\end{aligned}$$

Problem H4.5 continued on next page...

Thus a valid c is $c = 2$.

Part Three

$$\begin{aligned} n^2 - n + 1 &= \\ &\leq n^2 \\ &\leq c \cdot n^2/2 \end{aligned}$$

Thus a valid c is $c = 2$.

Problem H4.6

Let $\Sigma = \{0, 1\}$. Construct a DFA A that recognizes the language that consists of all binary numbers that can be divided by 5.

Let the state q_k indicate the remainder of k divided by 5. For example, the remainder of 2 would correlate to state q_2 because $7 \mod 5 = 2$.

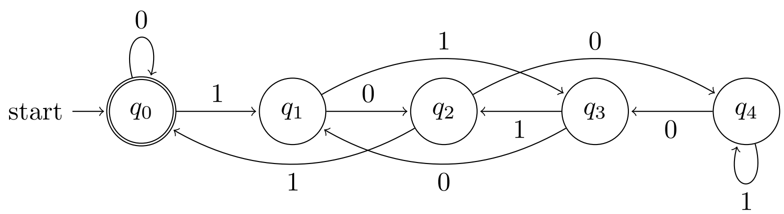

Figure 1: DFA, A, this is really beautiful, ya know?

Justification

Take a given binary number, x. Since there are only two inputs to our state machine, x can either become $x0$ or $x1$. When a 0 comes into the state machine, it is the same as taking the binary number and multiplying it by two. When a 1 comes into the machine, it is the same as multipying by two and adding one.

Using this knowledge, we can construct a transition table that tell us where to go:

	$x \mod 5 = 0$	$x \mod 5 = 1$	$x \mod 5 = 2$	$x \mod 5 = 3$	$x \mod 5 = 4$
$x0$	0	2	4	1	3
$x1$	1	3	0	2	4

Therefore on state q_0 or ($x \mod 5 = 0$), a transition line should go to state q_0 for the input 0 and a line should go to state q_1 for input 1. Continuing this gives us the Figure 1.

Problem H4.7

Write part of **Quick-Sort(**$list, start, end$**)**

function QUICK-SORT($list, start, end$)
  if $start \geq end$ **then**
    return
  end if
  $mid \leftarrow$ PARTITION($list, start, end$)
  QUICK-SORT($list, start, mid - 1$)
  QUICK-SORT($list, mid + 1, end$)
end function

Algorithm 1: Start of QuickSort

Problem H4.8

Suppose we would like to fit a straight line through the origin, i.e., $Y_i = \beta_1 x_i + e_i$ with $i = 1, \ldots, n$, $\mathrm{E}[e_i] = 0$, and $\mathrm{Var}[e_i] = \sigma_e^2$ and $\mathrm{Cov}[e_i, e_j] = 0, \forall i \neq j$.

Part A

Find the least squares esimator for $\hat{\beta}_1$ for the slope β_1.

Solution

To find the least squares estimator, we should minimize our Residual Sum of Squares, RSS:

$$
\begin{aligned}
RSS &= \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 \\
&= \sum_{i=1}^{n} \left(Y_i - \hat{\beta}_1 x_i\right)^2
\end{aligned}
$$

By taking the partial derivative in respect to $\hat{\beta}_1$, we get:

$$
\frac{\partial}{\partial \hat{\beta}_1}(RSS) = -2 \sum_{i=1}^{n} x_i (Y_i - \hat{\beta}_1 x_i) = 0
$$

This gives us:

$$
\begin{aligned}
\sum_{i=1}^{n} x_i (Y_i - \hat{\beta}_1 x_i) &= \sum_{i=1}^{n} x_i Y_i - \sum_{i=1}^{n} \hat{\beta}_1 x_i^2 \\
&= \sum_{i=1}^{n} x_i Y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2
\end{aligned}
$$

Solving for $\hat{\beta}_1$ gives the final estimator for β_1:

$$
\hat{\beta}_1 = \frac{\sum x_i Y_i}{\sum x_i^2}
$$

Part B

Calculate the bias and the variance for the estimated slope $\hat{\beta}_1$.

Solution

For the bias, we need to calculate the expected value $\mathrm{E}[\hat{\beta}_1]$:

$$\begin{aligned}
\mathrm{E}[\hat{\beta}_1] &= \mathrm{E}\left[\frac{\sum x_i Y_i}{\sum x_i^2}\right] \\
&= \frac{\sum x_i \mathrm{E}[Y_i]}{\sum x_i^2} \\
&= \frac{\sum x_i (\beta_1 x_i)}{\sum x_i^2} \\
&= \frac{\sum x_i^2 \beta_1}{\sum x_i^2} \\
&= \beta_1 \frac{\sum x_i^2 \beta_1}{\sum x_i^2} \\
&= \beta_1
\end{aligned}$$

Thus since our estimator's expected value is β_1, we can conclude that the bias of our estimator is 0.

For the variance:

$$\begin{aligned}
\mathrm{Var}[\hat{\beta}_1] &= \mathrm{Var}\left[\frac{\sum x_i Y_i}{\sum x_i^2}\right] \\
&= \frac{\sum x_i^2}{\sum x_i^2 \sum x_i^2} \mathrm{Var}[Y_i] \\
&= \frac{\sum x_i^2}{\sum x_i^2 \sum x_i^2} \mathrm{Var}[Y_i] \\
&= \frac{1}{\sum x_i^2} \mathrm{Var}[Y_i] \\
&= \frac{1}{\sum x_i^2} \sigma^2 \\
&= \frac{\sigma^2}{\sum x_i^2}
\end{aligned}$$

Problem H4.9

Prove a polynomial of degree k, $a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n^1 + a_0 n^0$ is a member of $\Theta(n^k)$ where $a_k \ldots a_0$ are nonnegative constants.

Proof. To prove that $a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n^1 + a_0 n^0$, we must show the following:

$$\exists c_1 \exists c_2 \forall n \geq n_0,\ c_1 \cdot g(n) \leq f(n) \leq c_2 \cdot g(n)$$

For the first inequality, it is easy to see that it holds because no matter what the constants are, $n^k \leq a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n^1 + a_0 n^0$ even if $c_1 = 1$ and $n_0 = 1$. This is because $n^k \leq c_1 \cdot a_k n^k$ for any nonnegative constant, c_1 and a_k.

Taking the second inequality, we prove it in the following way. By summation, $\sum\limits_{i=0}^{k} a_i$ will give us a new constant, A. By taking this value of A, we can then do the following:

$$\begin{aligned}
a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n^1 + a_0 n^0 = & \\
& \leq (a_k + a_{k-1} \ldots a_1 + a_0) \cdot n^k \\
& = A \cdot n^k \\
& \leq c_2 \cdot n^k
\end{aligned}$$

where $n_0 = 1$ and $c_2 = A$. c_2 is just a constant. Thus the proof is complete. $\square$

Problem H4.18

Evaluate $\sum_{k=1}^{5} k^2$ and $\sum_{k=1}^{5}(k-1)^2$.

Problem H4.19

Find the derivative of $f(x) = x^4 + 3x^2 - 2$

Problem H4.6

Evaluate the integrals $\int_0^1 (1 - x^2)\mathrm{d}x$ and $\int_1^\infty \frac{1}{x^2}\mathrm{d}x$.