1 Problem description

- Let (V, E) be a graph with vertices V ($|V|$ finite) and directed edges $E \subset V \times V$ where $(x, y) \in E$ indicates a directed edge from x to y.
- Denote by $\mathbb{R}^V$ and $\mathbb{R}^E$ the real vector spaces of functions defined on vertices and edges. We equip them with the standard Euclidean inner products:
$$\langle f, g \rangle_V = \sum_{x \in V} f(x) \cdot g(x) \quad \text{for} \quad f, g \in \mathbb{R}^V,$$
$$\langle \phi, \psi \rangle_E = \sum_{e \in E} \phi(e) \cdot \psi(e) \quad \text{for} \quad \phi, \psi \in \mathbb{R}^E.$$
- A function $\omega \in \mathbb{R}^E$ can be interpreted as flow on the edges where for $(x, y) \in E$ the value $\omega((x, y))$ denotes the mass flow from vertex x to y. $\omega((x, y)) < 0$ will be interpreted as mass flowing from y to x.
- For a given vertex $x \in V$ the net amount of mass leaving x by the flow ω can be expressed as
$$(\operatorname{div} \omega)(x) = \sum_{y \in V : (x,y) \in E} \omega((x, y)) - \sum_{y \in V : (y,x) \in E} \omega((y, x))$$
The first term sums contributions from edges starting at x, the second term sums contributions from edges ending at x.
- We can evaluate this expression for every x, hence the operator div is a linear map from $\mathbb{R}^E$ to $\mathbb{R}^V$. $\operatorname{div} \omega$ is called the *divergence* of ω.
- Let $\mathcal{P}(V) = \{\mu \in \mathbb{R}^V_+ : \sum_{x \in V} \mu(x) = 1\}$ be the space of discrete probability densities over V. For two $\mu, \nu \in \mathcal{P}(V)$ a flow $\omega \in \mathbb{R}^E$ is said to be a flow from μ to ν if $\operatorname{div} \omega = \nu - \mu$. If the graph (V, E) is connected (ignoring the orientation of edges), such a flow always exists.
- Assume each edge has a positive length, prescribed by a vector $L \in \mathbb{R}^E_{++}$. The *cost* of a flow is defined to be
$$\mathcal{C}(\omega) = \sum_{e \in E} |\omega(e)| \cdot L(e)\,.$$
This is the amount of mass flowing along each edge multiplied by the edge length.
- For fixed $\mu, \nu \in \mathcal{P}(V)$ we will study the following optimization problem:
$$\inf \left\{ \mathcal{C}(\omega) \middle| \omega \in \mathbb{R}^E, \operatorname{div} \omega = \nu - \mu \right\} = \inf_{\omega \in \mathbb{R}^E} \mathcal{C}(\omega) + \theta(\operatorname{div} \omega) \tag{1}$$
where $\theta : \mathbb{R}^V \to \mathbb{R} \cup \{\infty\}$ is given by $\theta = \iota_{\{\nu - \mu\}}$. We assume that the graph is connected.

2 Preliminaries

(i) Check that $\mathcal{C}$ and θ are convex, lower semicontinuous and proper.

(ii) Find the Fenchel–Legendre conjugates $\mathcal{C}^*$ and θ^*.

(iii) Find the subdifferentials of $\mathcal{C}$, θ, $\mathcal{C}^*$ and θ^*.

(iv) Show that the optimization problem (1) has a minimizer.

3 Duality

We state a variant of the Fenchel–Rockafellar Theorem.

Theorem 3.1. Let X, Y be Hilbert spaces, $A : X \to Y$ linear and bounded. Let $F : X \to \mathbb{R} \cup \{\infty\}$, $G : Y \to \mathbb{R} \cup \{\infty\}$ be convex, lower semicontinuous and proper. Assume there is a point $x \in X$ such that $F(x) < \infty$, $G(Ax) < \infty$ and G is continuous at Ax. Then

$$\inf_{x \in X} \left[F(x) + G(Ax)\right] = -\min_{y \in Y} \left[F^*(-A^* y) + G^*(y)\right]. \tag{2}$$

(i) Check that the assumptions for the Fenchel–Rockafellar Theorem are *not* met when identifying $X = \mathbb{R}^E$, $Y = \mathbb{R}^V$, $F = \mathcal{C}$, $G = \theta$, $A = \mathrm{div}$.

(ii) Derive the adjoint operator of div.

(iii) Ignoring that the assumptions are not met, nevertheless state the dual problem to (1).

(iv) Show that the assumptions for the Fenchel–Roackafellar Theorem are met when one considers the dual of (1) as primal in (2).

(v) Show that the dual problem of (1) has a solution.

(vi) Find a sufficient and necessary condition for a pair $(x, y) \in X \times Y$ to be primal and dual optimizers of (2).

4 Metric

For given $\mu, \nu \in \mathcal{P}(V)$ denote by $D(\mu, \nu)$ the corresponding optimal value of (1).

(i) Show that D is non-negative, symmetric and satisfies the triangle inequality. *Hint:* From minimal flows for $D(\mu, \nu)$ and $D(\nu, \rho)$ try to construct a feasible flow for $D(\mu, \rho)$.

D is called '*earth mover's distance*' or Wasserstein-1 distance on $\mathcal{P}(V)$.

5 Optimization

We state a variant of a proximal primal dual algorithm.

Theorem 5.1. Consider the setup of Theorem 3.1. Assume primal and dual problem have solutions. For $\tau, \sigma \in \mathbb{R}_{++}$, $\tau\sigma < \|A\|^{-2}$ and $(x^{(0)}, y^{(0)}) \in X \times Y$ let

$$x^{(\ell+1)} = \mathrm{Prox}_{\tau F}(x^{(\ell)} - \tau A^* y^{(\ell)}), \tag{3a}$$
$$y^{(\ell+1)} = \mathrm{Prox}_{\sigma G^*}(y^{(\ell)} + \sigma A(2x^{(\ell+1)} - x^{(\ell)})). \tag{3b}$$

Then $x^{(\ell)} \rightharpoonup x$, $y^{(\ell)} \rightharpoonup y$ as $\ell \to \infty$ where (x, y) are a pair of primal and dual solutions.

(i) Show that fixed-points of the iteration (3) are precisely the pairs of primal and dual solutions to (2).

(ii) Analogous to the Moreau decomposition express $\mathrm{Prox}_{\gamma f}$ via $\mathrm{Prox}_{\eta f^*}$ for a suitable factor η.

(iii) Find $\mathrm{Prox}_{\tau\mathcal{C}}$, $\mathrm{Prox}_{\tau\mathcal{C}^*}$, $\mathrm{Prox}_{\sigma\theta}$, $\mathrm{Prox}_{\sigma\theta^*}$.

6 Projection

If we want to solve (1) with the Douglas–Rachford algorithm we need to be able to compute $\mathrm{Prox}_{\tau H}$ where $H(\omega) = \theta(\operatorname{div} \omega)$. We find that $\mathrm{Prox}_{\tau H} = P_S$ where $S = \{\omega \in \mathbb{R}^E : \operatorname{div} \omega = \nu - \mu\}$.
More generally, let X and Y be two real Hilbert spaces, let $A : X \to Y$ be linear and bounded. For fixed $y \in Y$ let $S = \{x \in X : Ax = y\}$. Assume that $S \neq \emptyset$, i.e. y is in the image of A. We will determine how to compute the projection P_S.
The point $z = P_S x$ is the solution to

$$\min_{z \in X} \tfrac{1}{2}\|x - z\|^2 + \iota_{\{y\}}(Az) \tag{4}$$

(i) Use Theorem 2 to derive a dual problem to (4). Is (4) the primal or dual problem?

The problem dual to (4) will have the form

$$\inf_{z \in Y} \tfrac{1}{2}\|A^* z\|^2 + \langle z, b\rangle \tag{5}$$

for some $b \in Y$.

(ii) Find the solution of (5) for the case when AA^* is invertible.

(iii) Assume $\dim \ker A^* > 0$. Find a necessary and sufficient criterion for b such that (5) has a solution. What is the interpretation of this criterion in the problem (4)?

(iv) Use the relation 3(vi) to obtain the solution to (4) from the solution of (5).

7 Analysis of a convex function

In a slightly more complicated flow optimization problem the following function plays a central role:

$$\phi : \mathbb{R}^2 \to \mathbb{R} \cup \{\infty\}, \qquad (a, b) \mapsto \begin{cases} \frac{|b|^2}{a} & \text{if } a > 0, \\ 0 & \text{if } a = b = 0, \\ +\infty & \text{else.} \end{cases} \tag{6}$$

(i) Determine the sublevel sets of ϕ. Use this to show that ϕ is lower semicontinuous.

(ii) Construct a converging sequence $(a_n, b_n)_n$, $(a_n, b_n) \to (a, b) \in \mathbb{R}^2$ such that $\phi(a_n, b_n) < \infty$ and $\liminf_{n\to\infty} \phi(a_n, b_n) > \phi(a, b)$.

(iii) Find ϕ^*. Show that $\phi^{**} = \phi$.

(iv) Find the subdifferential of ϕ. *Hint:* Distinguish the cases $a > 0$ and $(a, b) = (0, 0)$. Note that ϕ is positively 1-homogeneous.

(v) Find $\mathrm{Prox}_{\tau\phi}$ for $\tau > 0$. *Note:* for some cases computing $\mathrm{Prox}_{\tau\phi}$ requires finding the root of a polynomial of degree three. It suffices to state how the root of the polynomial relates to $\mathrm{Prox}_{\tau\phi}$. An explicit formula for the root need *not* be given.

(vi) Since ϕ is positively 1-homogeneous, $\phi^* = \iota_B$ for some set $B \subset \mathbb{R}^2$. Find the normal cone of B. Use this to find an equation for the projection P_B onto B (Similar to above, the equation does not need to be solved explicitly.)

(vii) Verify the Moreau decomposition for Prox_ϕ and $\mathrm{Prox}_{\phi^*} = P_B$.