1. Exam program

The exam will take place on Tuesday May 05, 15-17h, in the Ecole Polytechnique.

(1) Ch. 1, convex functions: separation of convex sets (section 1.2), conjugacy, biconjugate, subdifferential, local continuity of locally upper bounded convex functions, cor. 1.58, tangent and convex cones in the sense of convex analysis.

(2) Ch. 1, duality (section 1): weak duality (section 1.3, sup inf ≤ inf sup), linear duality (section 1.4).

(3) Ch. 1, duality (section 2): perturbation duality, link with Lagrangian duality (rem. 1.70), strong duality (sec. 2.1.3), composite functions and Fenchel duality (sec. 2.1.4-5), subdifferential calculus (sec. 2.2).

(4) Integration: ch. 2, sec. 1-2 (except sec. 2.7).
2. Lesson 1 (Jan. 20, 2015): Convex functions

(1) Motivation: electricity production.
(2) Setting: Banach space $X$ (although the separation theorems are valid in normed vector spaces). Here $f$ is a function $X \to \mathbb{R}$, not necessarily convex.
(4) Application: a convex function $f$, continuous at $\bar{x}$, has a subgradient at this point. Hint: separation of $(\bar{x}, f(\bar{x}))$ and of the epigraph of $f$.
(5) Continuity: a convex function, uniformly bounded from above near $\bar{x}$, is Lipschitz near $\bar{x}$.
(6) Fenchel conjugates: Fenchel-Young inequality (FYI) $f(x) + f^*(x^*) \geq \langle x^*, x \rangle$.
   Equality iff $x^* \in \partial f(x)$.
   Biconjugate $f^{**}$ equal to the supremum of affine minorants of $f$.
(7) Fenchel Moreau Rockafellar theorem (FMR): biconjugate either always equal to $-\infty$, or equal to the convex closure.
(8) Directional derivatives: end of section 1.5.3 starting with lemma 1.55.
(9) Convex cones. In section 1.5.4. Polarity for cones. Tangent and normal cones, their expression when the convex set is itself a cone.


(1) Review of previous lesson: Convex closure and the FMR theorem.
   A locally dominated convex function (finitely valued at the reference point $\bar{x}$) is locally Lipschitz. At $\bar{x}$ its directional derivatives are the support function of its subdifferential.
(2) General duality: section 2.1.
   Abstract setting and strong duality (sections 2.1.1 and 2.1.3).
   Composite functions and Fenchel duality (sections 2.1.4 and 2.1.5).

(3) Subdifferential calculus: sections 2.2.
(4) Non convex duality: section 2.5.
(5) Additional material: see section 10.


(1) Review of previous lesson: Fenchel dual; direct derivation with the duality Lagrangian.
(2) Integration theory, ch. 2, especially the Egoroff theorem, Doob-Dynkin lemma, Fatou lemma.
(3) Ch. 2 again: Conjugates of expectations I: Carathéodory integrands, Castaing representation. Conjugate of a integral functional in the case of a Carathéodory integrands, in $L^s$ with $s \in [1, \infty)$.

5. LESSON 4 (FEB. 17, 2015): SUBGRADIENT OF EXPECTATIONS II  

(1) Exercices (see below).  
(2) Normal convex integrands (ch. 2, section 2.3).  
(3) Structure of elements in the dual of $L^\infty$. Subgradients of integrands in $L^\infty$ (ch. 2, section 2.4 et 2.5).

6. LESSON 5 (MARCH 3, 2015)  

(1) Conditional expectation: ch.5, section 1.  
(2) Dynamic stochastic programming I: ch.5, section 2.

7. LESSON 6 (MARCH 17, 2015)  

(1) Dynamic stochastic programming II: ch.5, end of section 2.  
(2) Dynamic programming for controlled Markov chains: ch. 6  
   (a) Finite horizon: section 1.2.  
   (b) Discounted problems: section 1.3.

8. LESSON 7 (MARCH 24, 2015)  

(1) Ergodic Markov chains: ch. 6, section 3.  
(2) Chapter 7:  
   Dynamic programming: continuous states  
   (a) Dynamic programming for deterministic and stochastic problems.  
   (b) Discretization: link with Markov chains.

9. LESSON 8 (LAST LESSON: MARCH 31, 2015)  

(1) Scalar unit commitment, exercice 10.9 below.  
(2) Probability constraints: exercices 10.11, 10.12 and 10.13 below.

10. ADDITIONAL MATERIAL

10.1. Convex analysis and duality. Let $\Omega$ be a bounded open subset of $\mathbb{R}^n$ wit smooth boundary. Let $X := H^1_0(\Omega)$ denote the Sobolev space of measurable function over $\Omega$ with square integrable derivatives, with norm such that $\|u\|^2_{H^1_0(\Omega)} = \int_\Omega |\nabla u(x)|^2 dx$, and dual denoted by $X^* = H^{-1}(\Omega)$. We identify $Y = L^2(\Omega)$ (but not $X$) with its dual.

Exercice 10.1. i) Given $u \in X$, show that $u^* := -\Delta u$ is the unique element of $X^*$ such that  

\[ (u^*, v)_X = (u, v)_X, \quad \text{for all } v \in X. \]
ii) Deduce that \( \varphi(u) := \frac{1}{2} \|u\|^2_{H^1_0(\Omega)} \) satisfies

\[
D \varphi(u) = -\Delta u \quad \text{(equality in } X^*), \text{ for all } u \in X.
\]

**Example 10.2 (Denoising with the } H^1_0(\Omega) \text{ norm).** We have a noisy measure \( u_d \in L^2(\Omega) \). The idea is to eliminate the noise by solving the problem, for some \( \varepsilon > 0 \) (regularizing parameter):

\[
\text{(3) } \min_{u \in X} J(u) := \frac{1}{2} \int_\Omega (u(x) - u_d(x))^2 \, dx + \frac{1}{2\varepsilon} \int \|\nabla u(x)\|^2 \, dx.
\]

We apply the Fenchel duality (with here \( x^* = 0 \) and \( y = 0 \)) where the operator \( A \) is the injection \( X \hookrightarrow Y \), so that \( A^\top \) is the restriction \( Y = Y^* \hookrightarrow X^* \), and

\[
\text{(4) } f(u) := \frac{1}{2\varepsilon} \|u\|^2_{H^1_0(\Omega)}; \quad F(y) := \frac{1}{2} \|y - u_d\|_Y^2.
\]

One can check that

\[
\text{(5) } f^*(u^*) = \frac{1}{2} \varepsilon^{-1} \|u^*\|^2_{H^{-1}(\Omega)}; \quad F^*(y^*) = (y^*, u_d)_Y + \frac{1}{2} \|y^*\|_Y^2.
\]

So the dual problem is

\[
\text{(6) } \max_{y^* \in Y} -\frac{1}{2} \varepsilon^{-1} \|y^*\|^2_{H^{-1}(\Omega)} - (y^*, u_d)_Y + \frac{1}{2} \|y^*\|_Y^2.
\]

The optimality condition reads, using the previous exercise:

\[
\text{(7) } y^* \in \partial F(u) = u - u_d; \quad -y^* \in \partial f(u) = -\varepsilon \Delta u.
\]

Therefore (as expected)

\[
\text{(8) } -\varepsilon \Delta u + u = u_d \quad \text{in } H^{-1}(\Omega).
\]

**Example 10.3 (Denoising with the } L^1 \text{ norm).** We have a black and white image with \( p \times q \) pixel, so that an image can be identified with a \( p \times q \) matrix \( u^d \). Its discrete gradient has values

\[
\text{(9) } (\nabla u)_{ij} := (u_{i+1,j} - u_{ij}, u_{i,j+1} - u_{ij}), \quad i = 1, \ldots, p - 1, \quad j = 1, \ldots, q - 1.
\]

In order to keep useful sharp variations (contours) one idea is to denoise by solving the problem

\[
\text{(10) } \min_u \frac{1}{2} \|u - u^d\|^2 + \varepsilon \|\nabla u\|_1.
\]

Here \( \| \cdot \| \) denotes as usual the Euclidean norm (since \( u \) is actually a matrix this is more precisely the Frobenius norm) and \( \| \cdot \|_1 \) denotes the discrete \( \ell^1 \) norm i.e. denoting by \( y_{i,j,1}, y_{i,j,2} \) the two components of \( y_{i,j} \):

\[
\text{(11) } \|y\|_1 := \sum_{0<i<p, 0<j<q} (|y_{i,j,1}| + |y_{i,j,2}|).
\]

This enters in the framework of Fenchel duality with

\[
\text{(12) } f(u) := \frac{1}{2} \|u - u^d\|^2; \quad Au := \nabla u; \quad F(y) := \varepsilon \|y\|_1.
\]

We have that

\[
\text{(13) } f^*(u^*) = (u^*, u^d) + \frac{1}{2} \|u^*\|^2; \quad F^*(y^*) = I_B(y^*/\varepsilon),
\]
where \( \bar{B} \) is the unit ball of \( L^\infty \). We denote \( A^\top \) by \( - \text{div} \) since it can be interpreted as a discrete version of the divergence operator. The Fenchel dual is

\[
\begin{align*}
\text{Max} & \quad - (\text{div} \, y^*, \, u^d) - \frac{1}{2} |\text{div} \, y^*|^2.
\end{align*}
\]

The optimality conditions read

\[
\begin{align*}
\text{div} \, y^* &= u - u^d; \quad y^*/\varepsilon \in \partial \|\cdot\|_1(\nabla u).
\end{align*}
\]

Since \( y \mapsto \|y\|_1 \) is positively homogeneous, setting \( \bar{y}^* := y^*/\varepsilon \), this is equivalent to

\[
\begin{align*}
\|\bar{y}^*\|_\infty &\leq 1; \quad \langle y^*, \nabla u \rangle = \|\nabla u\|_1; \quad u - u^d - \varepsilon \text{div} \, \bar{y}^* = 0.
\end{align*}
\]

The primal problem has a unique solution (coercive, strictly convex cost), and satisfies the stability condition. The dual problem has therefore a nonempty and bounded set of solution.

**Example 10.4 (A strange example).** Consider the reverse entropy function:

\[
\begin{align*}
\hat{H}(x) &= x \log x \text{ if } x > 0, \quad \hat{H}(0) = 0, \quad \text{and } \hat{H}(x) = +\infty \text{ if } x < 0.
\end{align*}
\]

This is a l.s.c. convex function, with domain \( \mathbb{R}_+ \). Consider the problem

\[
\begin{align*}
\text{Min} & \quad x \quad \text{ s.t. } \hat{H}(x) \leq 0.
\end{align*}
\]

It obviously has the unique solution \( \bar{x} = 0 \), and the stability condition holds (with here \( y = 0 \)). The Lagrangian of the problem is \( L(x, \lambda) := x + \lambda \hat{H}(x) \), and the dual problem is

\[
\begin{align*}
\text{Max} & \quad \delta(\lambda),
\end{align*}
\]

where \( \delta(\lambda) := \inf_{\lambda} L(x, \lambda) \). By the duality theory, the dual problem has a bounded and nonempty set of solutions and the primal and dual value are equal, i.e., \( \lambda \) is a dual solution iff \( \delta(\lambda) = 0 \), with infimum in the Lagrangian attained at 0. Now if \( \lambda > 0 \), the infimum is attained at a positive point. So, the unique dual solution is \( \bar{\lambda} = 0 \) and the optimality condition reads

\[
\begin{align*}
0 &\in \arg\min_{x \in \mathbb{R}} \left( x + 0 \times \hat{H}(\bar{\lambda}) \right).
\end{align*}
\]

This indeed holds if we interpret correctly the product \( 0 \times \hat{H}(\lambda) \) as being equal to \( +\infty \) whenever \( \hat{H}(\lambda) = +\infty \), see section 1.1.2 of the lecture notes.

10.2. Expectation minimization.

**Example 10.5 (Constrained entropy maximization).** Let \( \Omega \) be a measurable subset of \( \mathbb{R}^n \) with finite measure. Lebesgue measure. Consider the set of measurable, a.e. positive functions in \( X := L^1(\Omega) \):

\[
\begin{align*}
X_+ &:= \{ u \in L^1(\Omega); \quad u(\omega) \geq 0 \ \text{a.e.} \}.
\end{align*}
\]

We have observations

\[
\begin{align*}
\int_{\Omega} a_i(\omega) u(\omega) d\omega = b_i, \quad i = 1, \ldots, N,
\end{align*}
\]
where each $a_i$ belongs to $X^* = L^\infty(\Omega)$ and $b \in \mathbb{R}^N$ is a noisy measurement, so that the available information is that $b \in K$, where $K$ is a closed convex subset of $\mathbb{R}^N$. We define

$$H(u) := \int_\Omega \hat{H}(u(\omega))d\omega.$$  

We have in view cases when $u$ is a density probability and so we assume that $a_1(\omega) = 1$. In the crystallographic applications that we have in mind, $u(\omega)$ is the density probability for atoms to be at position $\omega$ and the observations correspond to the computation of Fourier modes, see [1]. The problem to be considered is

$$\text{Min}_{u \in X} H(u); \quad Au \in K,$$

where $(Au)_i := \int_\Omega a_i(\omega)u(\omega)d\omega$. Set

$$\hat{H}^\lambda(\omega, v) := \hat{H}(v) + \sum_{i=1}^N \lambda_i a_i(\omega) \cdot v.$$  

Let $a(\omega) := (a_1(\omega), \ldots, a_N(\omega))^\top$. Observe that

$$\inf_v \hat{H}^\lambda(\omega, v) = -\hat{H}^*(a(\omega) \cdot \lambda).$$

The Lagrangian function is

$$L(u, \lambda) := H(u) + \lambda^\top Au = \int_\Omega \hat{H}^\lambda(\omega, u(\omega))d\omega.$$  

The integrand is normal convex. Therefore, the dual cost satisfies

$$\delta(\lambda) = \inf_{u \in X} \int_\Omega \hat{H}^\lambda(\omega, u(\omega))d\omega - \sigma_K(\lambda) = -\int_\Omega \hat{H}^*(-a(\omega) \cdot \lambda)d\omega - \sigma_K(\lambda).$$

We assume that the stability condition holds. Since $\text{dom} H = X_+$, this condition can be expressed as

$$0 \in \text{int}(K - A(X_+)).$$

Since $\hat{H}(v) \geq -c$ for some $c > 0$ we have that the infimum of $H$ over the feasible set is not less than $-c|\Omega|$. By [29], the primal problem has a finite value, which is equal to the dual one, and the dual problem has at least a solution $\lambda$.

An element of $X_+$ is a primal solution iff it satisfies the optimality condition

$$\hat{H}(u(\omega)) + \hat{H}^*(-a(\omega) \cdot \lambda) = -(a(\omega) \cdot \lambda)u(\omega) \quad \text{a.e.}$$

Since $\hat{H}$ is strictly convex, there is a unique primal solution $\bar{u}$ that is determined by the above relation. Indeed we have that $D\hat{H}(v) = 1 + \log v = z$, iff $v = e^{z-1}$ and so

$$\bar{u}(\omega) = e^{-a(\omega) \cdot \lambda - 1}.$$  

It follows that

$$\hat{H}(\bar{u}(\omega)) = -(a(\omega) \cdot \lambda + 1)e^{-a(\omega) \cdot \lambda - 1}.$$
so that the dual cost is
\[(33) \quad \delta(\lambda) = -\int_{\Omega} e^{-a(\omega)\lambda - 1} d\omega - \sigma K(\lambda).\]

**Example 10.6.** Consider the particular case of the previous example when \(N = 1\), and the constraint to have a probability density, i.e. \(a(\omega) = 1\) a.e. and \(K = \{1\}\). Then the dual cost is \(-|\Omega|e^{-\lambda - 1} - \lambda\), which attains its maximum when \(|\Omega|e^{-\lambda - 1} = 1\), i.e., \(\lambda = \log |\Omega| - 1\); the optimal density is \(u = e^{-\lambda - 1} = 1/|\Omega|\) as expected.

**Example 10.7 (Phase transition models, see [2]).** Let \(f : \mathbb{R} \to \mathbb{R}, f(u) := u(1 - u)\), and let \(\Omega\) be a measurable subset of \(\mathbb{R}^n\). We choose the function space \(X := L^p(\Omega), p \in [1, \infty)\). For \(u \in X\), set \(F(u) := \int_{\Omega} f(u(\omega))d\mu(\omega)\), where \(d\mu\) is the Lebesgue measure. Consider the problem of minimizing \(F(u)\) with the constraints \(u(\omega) \in U\) a.e., \(U := [0, 1]\), and \(\int_{\Omega} u(\omega)d\mu(\omega) = a, a \in (0, \text{meas}(\Omega))\).

Given \(\lambda \in \mathbb{R}\), the Lagrangian of this problem is
\[(34) \quad L(u, \lambda) := F_U(u) + \lambda \left(\int_{\Omega} u(\omega)d\mu(\omega) - a\right) = \int_{\Omega} (f(u(\omega)) + \lambda u(\omega))d\mu(\omega) - \lambda a.
\]

The dual cost function is therefore
\[(35) \quad \delta(\lambda) = -F_U^*(-\lambda) - \lambda a = -\int_{\Omega} f_U^*(-\lambda)d\mu(\omega) - \lambda a.
\]

We compute
\[(36) \quad f_U^*(z) := \sup_{u \in U} uz - u(1 - u).
\]

Since \(u(1 - u)\) is concave the supremum is attained at 0 if \(z \leq 0\) and at 1 otherwise, and so,
\[(37) \quad f_U^*(z) = \begin{cases} 0 & \text{if } z \leq 0, \\ z & \text{if } z \geq 0. \end{cases}
\]

In other words \(f_U^*(z) = \max(0, z) = z_+\). So
\[(38) \quad \delta(\lambda) = -\int_{\Omega} (-\lambda)_+d\mu(\omega) - \lambda a = \begin{cases} \lambda(\text{meas}(\Omega) - a) & \text{if } \lambda \leq 0, \\ -\lambda a & \text{if } \lambda \geq 0. \end{cases}
\]

Clearly it attains its maximum at \(\bar{\lambda} = 0\), and so, the primal and dual values are equal, although the problem is nonconvex.

**Example 10.8.** This example illustrates how singular multipliers occur in optimality systems. Consider the problem
\[(39) \quad \min_{x \in \mathbb{R}} x; \quad x + 1/(k + 1) \geq 0, \quad k = 0, 1, \ldots.
\]

We choose \(\ell^\infty\) (the space of bounded sequences) as constraint space and denote by \(1\) and \(b\), respectively, the sequences of generic term \(1/(k + 1)\).
\[(40) \quad \min_{x \in \mathbb{R}} x; \quad x1 + b \geq 0, \quad k = 0, 1, \ldots
\]
where we have used the natural order relation for sequences. Let $K = \ell^\infty_+$ be the convex cone of elements of $\ell^\infty$ with nonnegative elements and let $A : \mathbb{R} \to \ell^\infty$, $Ax := x1$. The constraint can be written as $Ax + b \in K$. The duality Lagrangian is

$$x + \langle \lambda, x1 + b \rangle - \sigma_k(\lambda) = \langle \lambda, b \rangle + x(1 + \langle \lambda, 1 \rangle) - \sigma_k(\lambda).$$

The dual cone to $K$ is

$$K^- := \{ \lambda \in (\ell^\infty)^*; \langle \lambda, y \rangle \leq 0, \text{ for all } y \in \ell^\infty_+ \}.$$

So, the dual problem is

$$\text{Max}_{\lambda \in K^-} \langle \lambda, b \rangle; \quad 1 + \langle \lambda, 1 \rangle = 0.$$

The problem is convex and the stability condition obviously holds, and so, primal and dual values are equal. The optimality condition is, in view of the dual constraint:

$$0 = x - \langle \lambda, b \rangle = -\langle \lambda, x1 + b \rangle.$$

For any $y \in K$, $N_K(y) = K^- \cap y^\perp$ (see chapter 1), so that $K^- \cap (x1 + b)^\perp = N_K(x1 + b)$, the set of dual solutions (which is nonempty and bounded) is

$$\{ \lambda \in N_K(x1 + b); \quad 1 + \langle \lambda, 1 \rangle = 0 \}.$$

We now use the structure of elements of $(\ell^\infty)^*$. Any $\lambda \in (\ell^\infty)^*$ can be uniquely decomposed as $\lambda = \lambda^1 + \lambda^s$, where $\lambda^1 \in \ell^1$ and the singular part $\lambda^s$ depends only on the behavior at infinity.

For any $y \in K$, we have that

$$0 \geq \langle \lambda, y \rangle = \langle \lambda^1, y \rangle + \langle \lambda^s, y \rangle.$$

Taking, for $i \in \mathbb{N}$, $y = e_i$ (the sequence with all components equal to 0 except the $i$th equal to 1) we obtain that $\lambda^1 \in K^-$. Then let $y \in K$. Denote by $y^N$ the sequence whose $N$ first terms are zero, the other being equal to those of $y$. We have that $\langle \lambda^1, y^N \rangle = o(1)$ and $\langle \lambda^s, y^N \rangle = \langle \lambda^s, y \rangle$. Since $\lambda \in K^-$, we deduce that $\lambda^s \in K^-.$

Finally take $y = e_1/(k + 1)$. Then $x1 + \pm y \in K$, and therefore $0 \geq \langle \lambda, \pm y \rangle = \lambda^1_k/(k + 1)$, proving that $\lambda^1_k = 0$. and therefore $\lambda^1 = 0$. In view of the dual constraint, it follows that $\lambda^s \neq 0$.

10.3. Markov decision processe.

*Exercice* 10.9. [Scalar unit commitment] We have a demand $d^k$ belonging to $D := \{1, \ldots, n\}$. For $i \in D$, set $i_- := \max(0, i - 1)$ and $i_+ := \min(n, i + 1)$. We assume that $d^k$ is a Markov chain with transitions of probability $1/2$ from $i$ to $i_\pm$. We have a plant with state $e$, equal to 0 (off) or 1 (on). We pay $c_0 > 0$ if the plant is on to keep it on at the next time and $c_1 > c_0$ if the plant is off, to make it on at the next time. If the plant is on at time $k$, we have the reward $d^k$. So the total cost is 0 if keeping the plant off, $c_1$ when turning the plant on, $c_0 - d^k$ if maintaining it on, and $-d^k$ when turning off. The final value is zero. We denote the state by $(e, i)$.

We call 'monotonous' at time $k$ any policy such that, for some $i_k$ and $i'_k$ in $D$, one turns at time $k$ is $i \geq i'_k$, and keep the plant on if $i \geq i_k$. 


(1) Prove that the dynamic programming equation (DPP) is

\[
\begin{align*}
V_{0,i}^k &= \min \left( \frac{1}{2} (V_{0,i-}^{k+1} + V_{0,i+}^{k+1}), c_1 + \frac{1}{2} (V_{1,i-}^{k+1} + V_{1,i+}^{k+1}) \right), \\
V_{1,i}^k &= \min \left( \frac{1}{2} (V_{0,i-}^{k+1} + V_{0,i+}^{k+1}), c_0 + \frac{1}{2} (V_{1,i-}^{k+1} + V_{1,i+}^{k+1}) \right) - i.
\end{align*}
\]

(2) Show that the Bellman value is a nonincreasing function of \( e \) and \( d_k \) by two methods: (a) direct argument, (b) the DPP.

(3) Show that any optimal monotonous policy is such that \( i_k \leq i_k' \), and more generally that if, when the plant is on, an optimal decision is to turn off, then when the plant is off, the optimal decision is not to turn on.

(4) Compute \( V_{N-1}^{N-1} \).

Solution: \( V_{0,i}^{N-1} = 0, V_{1,i}^{N-1} = -i \).

(5) Compute \( V_{N-2}^{N-2} \) using the DPP and deduce that it is a monotonous policy.

Solution:

\[
\begin{align*}
V_{0,i}^{N-2} &= \min(0, c_1 - \frac{1}{2} (i_- + i_+)) \\
V_{1,i}^{N-2} &= \min(0, c_0 - \frac{1}{2} (i_- + i_+))
\end{align*}
\]

So an optimal policy is to turn on when \( c_1 \leq \frac{1}{2} (i_- + i_+) \), and to keep the plant on when \( c_0 \leq \frac{1}{2} (i_- + i_+) \), which is a monotonous policy.

(6) Show that any optimal policy is monotonous at time \( N-3 \).

Solution: we have that \( V_{0,i}^{N-1} - V_{1,i}^{N-1} \) is an increasing function of \( i \), and therefore also

\[
\Delta_i^k := \frac{1}{2} (V_{0,i-}^{k+1} + V_{0,i+}^{k+1}) - \frac{1}{2} (V_{1,i-}^{k+1} + V_{1,i+}^{k+1}).
\]

An optimal policy is to keep on when \( \Delta_i^k \geq c_0 \), and to turn on when \( \Delta_i^k \geq c_1 \). The result follows.

(7) Show that at least an optimal policy is monotonous.

Solution: ?

10.4. Dynamic programming.

Exercise 10.10. [Static unit commitment] We have \( n \) plants with switch-on cost \( a_i \) and variable production cost \( c_i(q_i) \) for producing an amount \( q_i \geq 0 \) of energy, with \( c_i \) convex function over \( \mathbb{R} \). We consider the problem of producing a total amount \( Q \) at minimum cost.

(1) Show that the problem can be formalized as follows:

\[
\min_{x,q} \sum_{i=1}^{n} x_i (a_i + c_i(q_i)); \quad \sum_{i=1}^{n} q_i \geq Q; \quad q_i \in [q_i^m, q_i^M], \quad i = 1, \ldots, n; \quad x \in \{0,1\}^n.
\]

(2) Show how to compute efficiently the dual function. What would be the cost if the \( c_i \) are quadratic ?

(3) Show how to solve efficiently the dual problem.

(4) Is a dual gap possible? Can we give an estimate of it?

(5) Can we make explicit the relaxed problem?
(6) Show how to solve the primal problem by a dynamic programming approach. Hint: as is classical for knapsack problems we decide at step $k$ how much energy will be produced by plant $k$, and the current state $\kappa$ is the sum of production of plants $1$ to $k - 1$. So the final cost is $0$ if $\kappa \geq Q$, and $+\infty$ if not, and the DP equation reads

$$
V^k(\kappa) = \min \left( V^{k+1}(\kappa), a_k + \min_{q_k \in [q_k^m, q_k^M]} (c_k(q_k) + V^{k+1}(\kappa + q_k)) \right).
$$

**Exercise 10.11. [Terminal probability constraints: duality]** We have dynamics and cost as in Ch.7, section 2, with the additional constraint

$$
P[y^N \in A, \text{for all } k] \geq \gamma \in [0, 1].
$$

(1) Show that we can reformulate the probabilistic constraint as the following expectation constraint

$$
\mathbb{E}1_{y^N \in A} \geq \gamma.
$$

(2) Is this constraint convex?

(3) Can we dualize it?

**Exercise 10.12. [Terminal probability constraints: martingale approach]** Inspired from Elie, Bouchard, Touzi, Pfeiffer ...

We consider the same problem as before.

(1) We add a state variable $z^k$ with constraint to be a submartingale, i.e.,

$$
z^k \leq \mathbb{E}_k z^{k+1}, \quad k = 0, \ldots, N - 1,
$$

and final state constraint

$$
z^N \leq 1_{y^N \in A}.
$$

Show that the probabilistic constraint holds if

$$
\mathbb{E}z^N \geq \gamma.
$$

(2) Show that the previous inequality holds in particular if $z^0 \geq \gamma$.

(3) Show that $z^k$ can be constrained to remain in $[0, 1]$.

(4) Design an algorithm based on this idea.


We have dynamics and cost as in Ch.7, section 2, with the additional constraint

$$
P[y^k \in A, \text{for all } k] \geq \gamma \in [0, 1].
$$

Consider the additional state $\pi$ with initial value $1$ and dynamics

$$
\pi^{k+1} = 1_{y^{k+1} \in A} \pi^k.
$$

Clearly

$$
\pi^N = \begin{cases} 
1 & \text{if } y^k \in A, \text{ for all } k, \\
0 & \text{otherwise}.
\end{cases}
$$
Therefore \( \mathbb{P}[y^k \in A, \text{for all } k] = \mathbb{E}\pi^N \), and so the probabilistic constraint can be reformulated as

\[
(60) \quad \mathbb{E}\pi^N \geq \gamma \in [0, 1].
\]

We can then dualize the constraint.

REFERENCES
