Exercise 1. Consider a probabilistic constraint given by

\[ X \triangleq \{ x \in \mathbb{R}^n : \mathbb{P}(g_i(x, \xi) \leq 0) \geq (1 - \alpha_i), \quad i = 1, \ldots, p \}, \]

where \( \alpha_i \in (0, 1) \) and \( g_i : \mathbb{R}^n \times \Omega \rightarrow \mathbb{R} \) are Carathéodory functions.

(i) Express \( \mathbb{P}(g_i(x, \xi) \leq 0) \) as an expectation of a suitable function.

(ii) Given a collection of samples \( \xi^1, \ldots, \xi^N \), using (i), provide an expression for a sample average approximation of \( X \), denoted by \( X_N \).

Definition 1 (Stochastic variational inequalities).

- Consider a random map \( F \), denoted by \( F(x, \omega) \), and defined as \( F : X \times \Omega \rightarrow \mathbb{R}^n \) where \( X \) is a fixed set \( \omega \in \Omega \). Associated with this map is a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and the associated expectation over this probability measure is denoted by \( \mathbb{E}[\cdot] \). The stochastic variational inequality problem, denoted by SVI\((X, F)\), requires finding a (deterministic) vector \( x \in X \) such that

\[ (y - x)^T \mathbb{E}[F(x, \omega)] \geq 0, \quad \forall y \in X. \quad (1) \]

- Note that the notation \( \mathbb{E}[F(x, \omega)] \) represents a vector in which the \( i \)-th component is given by \( \mathbb{E}[F_i(x, \omega)] \). In effect, SVI\((X, F)\) requires an \( x \in X \) such that \( -\mathbb{E}[F(x, \omega)] \in \mathcal{N}(x; X) \), while \( -F(x, \omega) \) may fail to belong to \( \mathcal{N}(x; X) \) for some \( \omega \in \Omega \).

- A mapping \( F(x) \) is strongly monotone over a set \( X \) if there exists an \( \eta > 0 \) such that

\[ (F(x) - F(y))^T (x - y) \geq \eta \| x - y \|^2, \quad \forall x, y \in X. \]

A mapping \( F(x) \) is Lipschitz continuous over a set \( X \) if there exists an \( L \) such that

\[ \| F(x) - F(y) \| \leq L \| x - y \|, \quad \forall x, y \in X. \]

The next Lemma is not to be proved.

Lemma 2. Let \( V_k \) be a sequence of non-negative random variables adapted to \( \sigma \)-algebra \( \mathcal{F}_k \) and such that \( a.s. \)

\[ \mathbb{E}[V_{k+1} | \mathcal{F}_k] \leq (1 - u_k)V_k + \beta_k \quad \text{for all } k \geq 0, \]

where \( 0 \leq u_k \leq 1, \beta_k \geq 0 \), and \( \sum_{k=0}^{\infty} u_k = \infty, \sum_{k=0}^{\infty} \beta_k < \infty, \lim_{k \to \infty}, \) and \( \frac{\beta_k}{u_k} = 0 \). Then, \( V_k \to 0 \) a.s. as \( k \to \infty \).
Exercise 3. Consider a sequence produced by the following update rule:

\[ x_{k+1} := \Pi_X (x_k - \gamma_k F(x_k, \omega_k)) , \]

where \( k \geq 0 \), \( \Pi_X(y) \) denotes the projection of \( y \) on the closed and convex set, \( \gamma_k \) denotes the steplength, and \( F(x_k, \omega_k) \) denotes the sampled map. Consider a single-valued continuous mapping \( F(x) \) that is Lipschitz continuous and strongly monotone over \( X \) with constants \( L \) and \( \eta \), respectively.

(i) By recalling that \( x^* \) satisfies

\[ x^* = \Pi_X(x^* - \gamma F(x^*)) , \]

for any positive scalar \( \gamma \) and \( w^k \triangleq F(x_k) - F(x_k, \omega_k) \), show that

\[ \|x_{k+1} - x^*\|^2 \leq \|x_k - x^* - \gamma_k F(x_k) + w_k - F(x^*)\|^2 . \]

(ii) Suppose \( \mathbb{E}[\|w^k\|^2 | \mathcal{F}_k] \leq \nu^2 \) almost-surely for \( k \geq 0 \). Then show that the following relation holds almost surely for all \( k \geq 0 \):

\[ \mathbb{E}[\|x_{k+1} - x^*\|^2 | \mathcal{F}_k] \leq (1 - 2\eta \gamma_k + L^2 \gamma_k^2) \|x_k - x^*\|^2 + \gamma_k^2 \nu^2 . \]

(iii) Suppose \( \gamma < \frac{2\eta}{L^2} \) then the scheme (2) produces a sequence \( \{x_k\} \) such that

\[ \lim_{k \to \infty} x_k = x^* , \]

where \( x^* \) is the unique solution of SVI(\( X, F \)). (use the super-martingale convergence Lemma).

Exercise 4. Consider the iterative proximal point scheme given by

\[ x^{k+1} = \Pi_X [x^k - \gamma_k (F(x^k) + \theta_k (x^k - x^{k-1}) + w^k)] , \]

where \( \gamma_k \) denotes the stepsize and \( \theta_k \) represents the weight on the centering parameter at the \( k \)th iteration.

(i) Show that the (IPP) scheme with \( w_k = 0 \) can be viewed as a standard projection scheme with a suitably defined \( x(\theta_k) \):

\[ x^{k+1} = \Pi_K [x^k(\theta) - \gamma_k F(x^k)] . \]

(ii) Show that

\[ \|x^{k+1} - x^*\|^2 \leq \|(x^k - x^*) - \gamma_k (F(x^k) - F(x^*) - \theta_k (x^k - x^{k-1}) - w^k)\|^2 . \]

(iii) By leveraging Lipschitz continuity and that \( \mathbb{E}[w^k | \mathcal{F}_k] = 0 \), show that the following bound can be derived:

\[
\mathbb{E}[\|x^{k+1} - x^*\|^2 | \mathcal{F}_k] \leq \|x^k - x^*\|^2 + \gamma_k^2 \|F(x^k) - F(x^*)\|^2 + \gamma_k^2 \mathbb{E}[\|w^k\|^2 | \mathcal{F}_k] \\
+ (\gamma_k \theta_k)^2 \|x^k - x^{k-1}\|^2 - 2\gamma_k (x^k - x^*)^T (F(x^k) - F(x^*)) \\
- 2\gamma_k \theta_k (x^k - x^*)^T (x^k - x^{k-1}) + 2\gamma_k^2 \theta_k (F(x^k) - F(x^*))^T (x^k - x^{k-1}) .
\]