Limites de suites
Problème de recherche : le réglage d'une intelligence artificielle

« Entraîner » une intelligence artificielle, c'est chercher les réglages qui rendent son erreur la plus petite possible. Pour cela, les ingénieurs utilisent presque toujours le même algorithme : la *descente de gradient*.
Dans une version simplifiée, l'erreur de l'IA dépend d'un seul réglage x et vaut

$$f(x) = 2(x-3)^2 + 1.$$

L'algorithme ne « voit » pas la courbe : il part d'un réglage $x_0 = 0$, puis, à chaque étape, il fait un petit pas dans le sens où l'erreur diminue, selon la règle

$$x_{n+1} = x_n - \eta\, f'(x_n),$$

où η (la lettre grecque « êta ») est un nombre strictement positif fixé à l'avance, appelé *taux d'apprentissage*.

La mission

1. Pour quelles valeurs de η la suite (x_n) converge-t-elle vers le meilleur réglage ?
2. Quelle valeur de η conseiller à l'équipe ? Que se passe-t-il pour les autres valeurs ?
3. Avec $\eta = 0{,}1$, à partir de quelle étape le réglage est-il à moins de 10^{-6} du meilleur réglage ?

Pour aller plus loin

Une IA plus réaliste a deux réglages x et y, et son erreur vaut $g(x\,;y) = x^2 + 10y^2$. L'algorithme part de $(x_0\,;y_0) = (1\,;1)$ et met à jour les deux réglages *avec le même taux* η :

$$x_{n+1} = x_n - \eta \times 2x_n \qquad \text{et} \qquad y_{n+1} = y_n - \eta \times 20y_n.$$

Pour quelles valeurs de η les deux réglages convergent-ils vers 0 ? Quelle valeur de η fait converger le plus vite *les deux à la fois* ?

Coups de pouce

Bloqué ? Des indices progressifs sont disponibles en ligne. N'en ouvrir qu'un à la fois, et seulement après avoir cherché !