Метод проекции градиента

Рассмотрим многомерную задачу локальной условной оптимизации

$$\min_{\mathbf{X} \in D \subset R^n} \Phi(\mathbf{X}) = \Phi(\mathbf{X}^*) = \Phi^*, \qquad (1)$$

где множество допустимых значений определяется только ограничениями типа неравенств

$$D = \{\mathbf{X} \mid g_j(\mathbf{X}) \geq 0, j \in [1, l] \qquad (2)$$

и целевая функция $\Phi(\mathbf{X})$ и ограничивающие функции $g_j(\mathbf{X}), j \in [1, l]$ являются непрерывными и дифференцируемыми функциями, а ограничивающие функции еще и выпуклы.

Проектирование точки на множество.

Идея *метода проекции градиента* состоит в том, что если на некоторой итерации точка

$$\mathbf{X}^{r+1} = \mathbf{X}^r + \lambda^r \mathbf{S}^r, \qquad (3)$$

полученная с помощью градиентного метода наискорейшего спуска (см. главу 7), оказывается вне множества допустимых значений D, то она возвращается на это множество. Возврат производится с помощью процедуры "*проекция точки на множество*". Напомним, что в формуле (3) λ^r — длина шага на r-ой итерации в направлении $\mathbf{S}^r$;

$$\mathbf{S}^r = -\frac{\nabla\Phi^r}{\|\nabla\Phi^r\|} -$$

единичный вектор направления антиградиента функции $\Phi(\mathbf{X})$ в точке $\mathbf{X}^r$, $\|*\|$ — некоторая векторная норма, например, евклидова.

Определение. Проекцией точки $\mathbf{X} \in R^n$ на замкнутое множество $D \subset R^n$ называется ближайшая к $\mathbf{X}$ точка $\hat{\mathbf{X}}$ множества D. Т.е. точка $\hat{\mathbf{X}} \in D$ называется проекцией точки $\mathbf{X} \in R^n$ на замкнутое множество $D \subset R^n$, если

$$\rho(\mathbf{X}, \hat{\mathbf{X}}) = \min_{\mathbf{Y} \in D} \rho(\mathbf{X}, \mathbf{Y}) \qquad (4)$$

где $\rho(\mathbf{A}, \mathbf{B})$ — расстояние между точками $\mathbf{A} \in R^n, \mathbf{B} \in R^n$ в некоторой метрике, например, $\rho(\mathbf{A}, \mathbf{B}) = \|\mathbf{A} - \mathbf{B}\|$ ●

Проекцию $\hat{\mathbf{X}}$ точки $\mathbf{X}$ на замкнутое множество D будем обозначать $P_D(\mathbf{X}) = \mathbf{X}$ (см. рис. 1). Очевидно, что $P_D(\mathbf{X}) = \mathbf{X}$, если $\mathbf{X} \in D$.

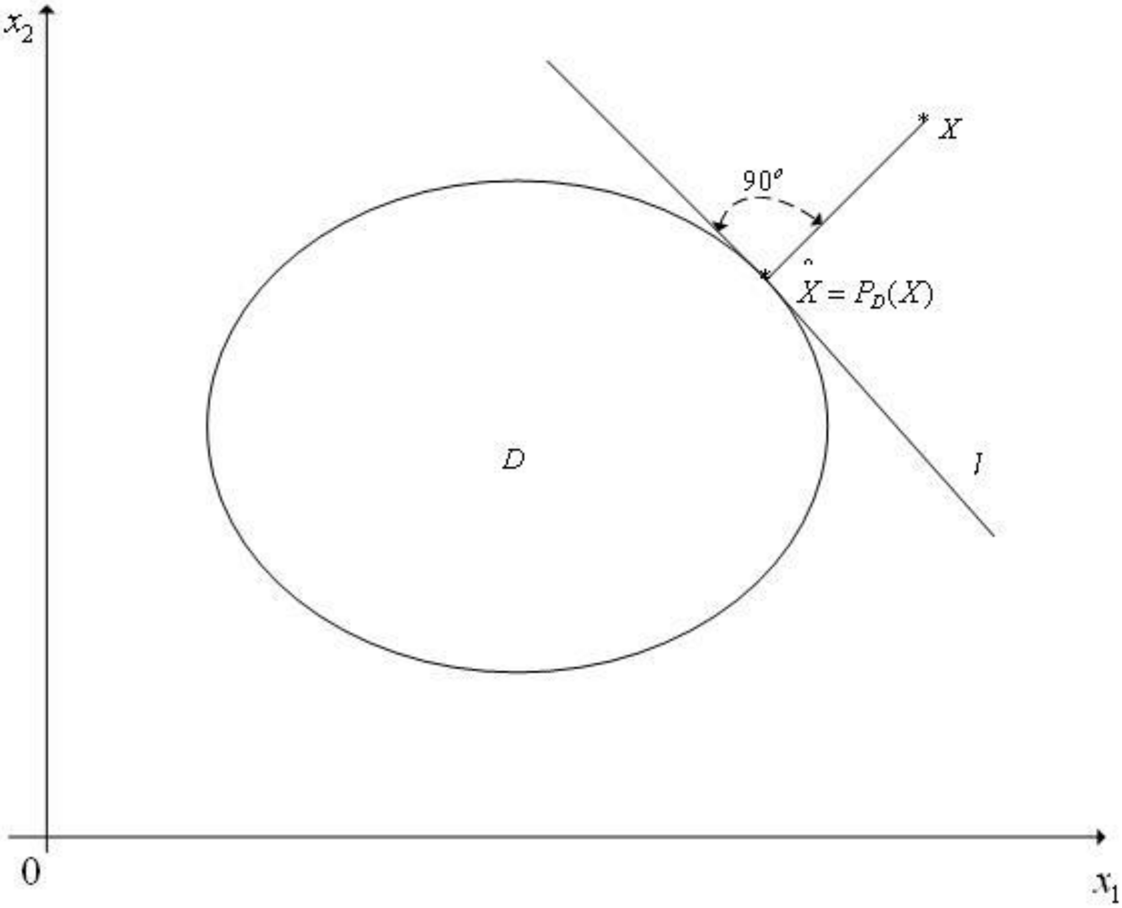

Рис. 1. К определению проекции точки на множество. Прямая l является касательной к границе области D в точке P_D(X)=X.

Можно показать, что если D — замкнутое выпуклое множество пространства R^n, то для любой точки $\mathbf{X} \in R^n$ существует единственная ее проекция на это множество.

Задача (4) поиска проекции точки на множество также является многомерной задачей условной оптимизации и ее решении может вызвать в общем случае значительные затруднения.

Задача (4) становится задачей квадратичного программирования, если множество D задается лишь линейными ограничениями типа неравенств и если функция $\rho(\mathbf{X}, \mathbf{Y})$ является квадратичной функцией $\mathbf{Y}$, например, если $\rho(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X} - \mathbf{Y}\|$.

Наибольший практический интерес представляет ситуация, когда множество D таково, что задача (4) может быть решена в явном виде. Приведем несколько наиболее практически важных примеров таких множеств.

Схема комбинации метода проекции градиента с методом дробления шага.

Метод проекции градиента может быть скомбинирован со многими градиентными методами (см. главу 7). Рассмотрим комбинацию метода проекции градиента с градиентным методом дробления шага.

Напомним, что в градиентном методе с дроблением шага величина шага λ^r находится из условия

$$\Phi(\mathbf{X}^r) - \Phi(\mathbf{X}^{r+1}) < 0.5\lambda^r \|\nabla\Phi^r\| \qquad (5)$$

Схема метода:

1. Задаем начальную точку $\mathbf{X}^0$, начальную величину шага λ^0 и коэффициент дробления шага $\nu \in (0, 1)$. Полагаем счетчик числа итераций $r = 0$.
2. По формуле (3) вычисляем координаты точки $\tilde{\mathbf{X}}^{r+1}$ и проекцию $P_D(\tilde{\mathbf{X}}^{r+1}) = \mathbf{X}^{r+1}$ этой точки на множество D.
3. Вычисляем величину $\Phi(\mathbf{X}^{r+1})$ — значение функции $\Phi(\mathbf{X})$ в точке $\mathbf{X}^{r+1}$.
4. Если условие дробления шага выполнено (см. параграф 7.1), то переходим к следующему пункту. Иначе – переходим к п.6.
5. Полагаем $\lambda^r = \nu\lambda^r$ и переходим к п.2.
6. Проверяем условие окончания поиска (см. ниже). Если условие окончания поиска выполнено, то полагаем $X^* \approx \mathbf{X}^{r+1}, \Phi^* \approx \Phi(\mathbf{X}^{r+1})$ и завершаем итерации. Иначе – полагаем $r = r+1$ переходим к п.2 •

В качестве критерия окончания поиска можно использоваться одно из стандартных условий окончания итераций

$\|\mathbf{X}^{r+1} - \mathbf{X}^r\| = \lambda^r \le \varepsilon_X$,

$|\Phi(\mathbf{X}^{r+1}) - \Phi(\mathbf{X}^r)| \le \varepsilon_\Phi$,

или условие $\|\nabla\Phi^r\| \le \varepsilon_\nabla$, где — константа, определяющая требуемую точность решения по градиенту функции $\Phi(\mathbf{X})$.

Комбинацию метода проекции градиента и градиентного метода с дроблением шага иллюстрирует рис. 2, на котором показан фрагмент линий уровня функции Химмельблау.

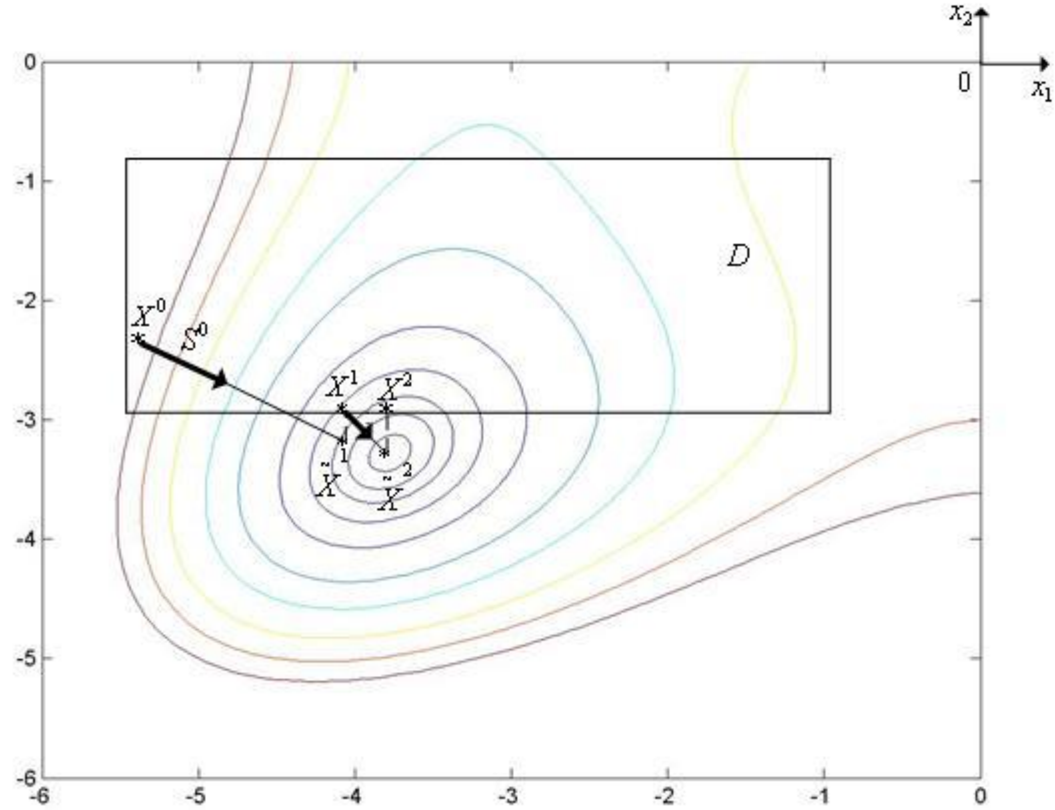

Рис. 2. Траектория поиска минимума функции Химмельблау комбинацией метода проекции градиента и градиентного метода с дроблением шага.

Известны модификации метода проекции градиента, ориентированные на решение задач условной оптимизации с ограничениями типа равенств.