Polynomial matrices
Quasi-linear extended GCD

Optional exercise for 2023-01-10

Let \( f \) and \( g \) be two univariate polynomials in \( \mathbb{K}[X] \) of respective degrees \( m \) and \( n \); in what follows we assume \( n > m > 0 \). Let \( h = \text{gcd}(f, g) \), and let \( \ell \) be the degree of \( h \), which satisfies \( \ell \leq m \). We also define \( \bar{f} = f/h \) and \( \bar{g} = g/h \), which are polynomials in \( \mathbb{K}[X] \) of respective degrees \( m - \ell \) and \( n - \ell \).

In the course, we have seen how to apply fast kernel basis computation (via a single approximant basis in that case) to solve the algorithmic problem of the GCD: given \( f \) and \( g \), compute \( h \). This was based on the following result: the left kernel of \( \begin{bmatrix} f \\ g \end{bmatrix} \) has rank 1 and a basis of it is given by \( \begin{bmatrix} -\bar{g} \\ \bar{f} \end{bmatrix} \). Then, the gcd can be obtained via a reduced kernel basis, which is itself computed efficiently via a basis of Hermite-Padé approximants for \( \begin{bmatrix} f' \\ g' \end{bmatrix} \) at order \( O(n) \), yielding the complexity \( O(M(n) \log(n)) \).

Here, we focus on the extended GCD, which further computes cofactors \( (u, v) \) such that \( uf + vg = h \). It is a classical result that there exists a unique pair of polynomials \((u, v)\) in \( \mathbb{K}[X]^2 \) such that

\[
\begin{cases}
  uf + vg = h, \\
  \deg(u) < n - \ell \quad \text{and} \quad \deg(v) < m - \ell.
\end{cases}
\]

(1)

Then, using this equation and the kernel basis given above, we obtain that for this pair \((u, v)\) in \( \mathbb{K}[X]^2 \),

\[
\begin{bmatrix}
  u \\
  -\bar{g}
\end{bmatrix}
\begin{bmatrix}
  f \\
  g
\end{bmatrix}
= \begin{bmatrix}
  uf + vg \\
  -\bar{g}f + fg
\end{bmatrix}
= \begin{bmatrix}
  h \\
  0
\end{bmatrix}.
\]

(2)

As seen previously in the course, each step of the extended Euclidean algorithm applied to \( f \) and \( g \) corresponds to an elementary transformation, represented by a \( 2 \times 2 \) polynomial matrix which is unimodular (i.e. its determinant is constant). The successive steps yield several such matrices: applying all steps successively corresponds to multiplying all these matrices together, which precisely gives the matrix \( \begin{bmatrix} u & v \\ -\bar{g} & \bar{f} \end{bmatrix} \). As a product of unimodular matrices, this matrix is unimodular as well. This can also be observed directly by considering its determinant \( uf + vg = (uf + vg)/h = 1 \). It follows that the constant term of this matrix is invertible, which will play a role below in reducedness properties:

\[
\det\left(\begin{bmatrix}
  u(0) \\
  -\bar{g}(0)
\end{bmatrix}
\begin{bmatrix}
  v(0) \\
  \bar{f}(0)
\end{bmatrix}\right) = u(0)f(0) + v(0)\bar{g}(0) = 1.
\]

The approach is to compute the matrix \( \begin{bmatrix} u & v \\ -\bar{g} & \bar{f} \end{bmatrix} \) by exploiting the fact that it satisfies Eq. (2). One obstacle is that the right-hand side of Eq. (2), involving the sought gcd \( h \), is unknown. Yet, we know that it has “small degree”, since by construction \( u \) and \( v \) are built to cancel high-degree terms in \( f \) and \( g \); indeed \( \deg(h) \leq \min(m, n) = m \), whereas the left-hand side of Eq. (2) involves \( uf \) and \( vg \) which both have degree up to \( m + n - \ell - 1 \). This problem of solving \( uf + vg = h \) for \((u, v, h)\) is akin to the problem of Euclidean division, solving \( f = qg + r \) for \((q, r)\) with \( r \) of “small degree”. Then, in a way similar to what was done earlier in the course for fast division with remainder, we will use reversals to “replace the unknown small degree polynomial \( h \) by a known power \( x^\delta \).

In what follows, for a polynomial \( p = p_0 + p_1 x + \cdots + p_\delta x^\delta \) of degree \( \delta \) and an integer \( N \geq \delta \), we denote the reversal of \( p \) with respect to \( N \) by

\[
\text{rev}(p, N) = x^N p(x^{-1}) = p_\delta x^{N-\delta} + \cdots + p_1 x^{N+1} + p_0 x^N.
\]
Should difficulties arise in the first question, it can be taken for granted.

1. For \((u, v) \in \mathbb{K}[X]^2\) such that Eq. (1), let
   \[
   R = \begin{bmatrix}
   \text{rev}(u, n - \ell - 1) & \text{rev}(v, m - \ell - 1) \\
   \text{rev}(\bar{g}, n - \ell) & \text{rev}(\bar{f}, m - \ell)
   \end{bmatrix} \in \mathbb{K}[X]^{2 \times 2}.
   \]
   Show that
   \[
   R \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix} = \begin{bmatrix} x^{m+n-2\ell-1} \text{rev}(h, \ell) \\ 0 \end{bmatrix} \qquad (3)
   \]

2. We define the shift \(s = (-n, -m) \in \mathbb{Z}^2\). Show that the matrix \(R\) defined above has \(s\)-row degree \((-\ell - 1, -\ell)\), and is an \(s\)-reduced basis of the module of Hermite-Padé approximants
   \[
   \mathcal{A} = \left\{ [p \ q] \in \mathbb{K}[X]^{1 \times 2} \bigg| \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix} = 0 \mod x^{m+n-2\ell-1} \right\}.
   \]

3. One issue is that, for algorithmic purposes, \(\ell\) is usually unknown. We will make use of an order \(d\) which only slightly overestimates \(n + m - 2\ell - 1\), and allows us to retrieve the extended gcd, thanks to the following consequence of the previous question.
   Show that, for any \(d \geq n + m - 2\ell - 1\), writing \(\delta = d - (n + m - 2\ell - 1)\), the matrix
   \[
   \begin{bmatrix}
   x^{\delta} & 0 \\
   0 & 1
   \end{bmatrix} R = \begin{bmatrix}
   x^{\delta} \text{rev}(u, n - \ell - 1) & x^{\delta} \text{rev}(v, m - \ell - 1) \\
   -\text{rev}(\bar{g}, n - \ell) & \text{rev}(\bar{f}, m - \ell)
   \end{bmatrix}
   \]
   is an \(s\)-reduced basis of the \(\mathbb{K}[X]\)-module of Hermite-Padé approximants
   \[
   \mathcal{A}_d = \left\{ [p \ q] \in \mathbb{K}[X]^{1 \times 2} \bigg| \begin{bmatrix} \text{rev}(f, m) \\ \text{rev}(g, n) \end{bmatrix} = 0 \mod x^d \right\}.
   \]
   Note that its \(s\)-row degree is \((-\ell + \delta - 1, -\ell)\).

**Hint:** for question 3, use the “product-based approach” for shifted reduced bases of modules, seen in the course. For example, you can use the particular case that was described for Hermite-Padé approximants in exercise 3 of last week’s exercises, but with two different orders \(m + n - 2\ell - 1\) and \(\delta\) for the two approximant bases (note that the sum of these orders is \(d\)):

(a) The first basis, for \(\mathcal{A}\), is the matrix \(R\) above.

(b) Verify that the residual matrix \(G\) is precisely \(\begin{bmatrix} \text{rev}(h, \ell) \\ 0 \end{bmatrix}\).

(c) Deduce that the basis for the second problem, for \(G\) at order \(\delta\), is \(\begin{bmatrix} x^\delta & 0 \\ 0 & 1 \end{bmatrix}\).

(d) Conclude by taking the product of these two bases.

**Remark/Perspective:** the algorithm suggested by question 3 may not give precisely the matrix \(R\), but any \(s\)-reduced basis of such approximants. Here are two ways to recover \((u, v, h)\):

- Ensure a sufficient degree discrepancy between the two rows for example by ensuring \(\delta \geq 3\) (which does not impact the asymptotic complexity). This allows us to distinguish the row giving \(u, v\) (the one with largest \(s\)-row degree) from the row with the kernel involving \(\bar{f}, \bar{g}\) (the one with smallest \(s\)-row degree).

- Use an algorithm which returns a basis with a stronger, canonical form called the \(s\)-Popov form.