### Motivation

Let \( q_1(x) = g_1^2(x)h(x), \) \( q_2(x) = g_2^2(x)h(x). \) We observe \( p_1, p_2, p_3 = q_1^2 + e_1 \), \( p_2 = q_2^2 + e_2. \)

Goal: recover the common divisor \( h \) and cofactors \( g_1 \) and \( g_2. \)

But, \( \gcd(p_1, p_2) = 1 \) for almost all \( e_1 \) and \( e_2. \)

⇒ we need a notion of approximate (greatest) common divisor.

### Problem statement

**Problem 1. (Approximate GCD with bounded degree)**

Given \( N \) polynomials \( p_1 \in P_n, \ldots, p_N \in P_n \) and degree \( d, \)

\[
\min_{p_i \in P_n} \sum_{k=1}^{N} \|p^k_1 - p^k_2\|_{w_i}^2 \quad \text{subject to} \quad \deg(\gcd(p_1^2, \ldots, p_N^2)) \geq d,
\]

where:

- \( P_n := \{p_n z^0 + \cdots + p_n z^n \mid p_n \in \mathbb{R}[z]\} \subset \mathbb{R}[z], \)
- \( \|p_2^0\|_{w_i}^2 = \sum_{w \in [0, \infty]} ^{w_{i+1}} \|w\| \quad \text{and} \quad \|w\| = \|w\|_{\infty} \equiv \text{max} \|w\| \)
- \( \|w\|_{\infty} = 0 \equiv \text{missing coefficients} (p_k = \hat{p}_k), \)
- \( \|w\| = \infty \equiv \text{fixed coefficients} (p_k = \hat{p}_k), \)

Note: Problem 1 is non-convex.

### Parameterizations of the problem

- **Image representation**
  \( \min_{\hat{g}, x \in X, \hat{h} \in h} \|\hat{g} - g^2\|_{\nu} \quad \text{(IM)} \)

  — nonlinear least-squares problem in \( \hat{g}, \hat{h}. \)

- **Kernel representation**
  \( \deg(\gcd(p_1, \ldots, p_N)) \geq d \iff \hat{g}(p) \quad \text{is rank-deficient,} \)
  \where \( p := (\hat{g}, \ldots, \hat{g}) \) \( \text{and} \hat{g}(p) \) \( \text{is a Sylvester-like matrix.} \)

  Example: \( \hat{g}(a, b) = \begin{bmatrix} a_0 & a_1 & \ldots & a_n \\ b_0 & b_1 & \ldots & b_n \\ \vdots & \vdots & \ddots & \vdots \\ n_0 & n_1 & \ldots & n_n \end{bmatrix} \in \mathbb{R}^{(m+n) \times (m+n)}. \)

  ⇒ Problem 1 is a structured low-rank approximation problem:

\[
\min_{\hat{g} \in P_N} \sum_{k=1}^{N} \|p^k_1 - p^k_2\|_{w_i}^2 \quad \text{subject to} \quad \hat{g}(p) \quad \text{is rank-deficient.}
\]

### Variable projection principle

Let \( F: \mathbb{R}^{m+n} \rightarrow \mathbb{R}^M \) such that \( F(x, y) = A_2 y + b, \forall x \in \mathbb{X}, \mathbb{X} \subseteq \mathbb{R}^m. \) Then the following nonlinear least squares problem

\[
\min_{x \in \mathbb{X}, y \in \mathbb{R}} \|F(x, y)\|_2^2 \quad \text{(NLS)}
\]

is equivalent to

\[
\min_{x \in \mathbb{X}} \min_{y \in \mathbb{R}} \|F(x, y)\|_2^2 \quad \text{(INNER)}
\]

\[
\min_{x \in \mathbb{X}} \min_{y \in \mathbb{R}} \min_{y \in \mathbb{R}} \|F(x, y)\|_2^2 \quad \text{(OUTER)}
\]

### Advantages:

- **(INNER)** is a linear least squares problem, has a closed-form solution.
- the number of variables in nonlinear optimization is reduced from \( m+n \) to \( m. \)
- we can evaluate derivatives of \( f(x) \) and use general-purpose optimization methods for (OUTER).
- (OUTER) is also a nonlinear least squares problem \( f(x) = \|F(x)\|_2^2 \text{ for some } F: \mathbb{R}^m \rightarrow \mathbb{R}^4, \) and we can use specialized methods (e.g. Levenberg-Marquardt).

### Application to different parameterizations

- **Image representation**
  The cost function in (IM) has the form
  \[
  \|\hat{g} - \hat{h}\|_2^2
  \]
  where \( \hat{g} := (\hat{g}_1, \ldots, \hat{g}_N) \) and \( F \) is biaffine in \( \hat{g} \) and \( \hat{h}. \)

  ⇒ we can apply the variable projection principle.

  Two options:

  ∗ take \( x = \hat{h}, y = \hat{g}; \) \( \mathbb{X} := \{x : \|x\|_2 = 1\} \) (denoted by \( \text{IM}_{\hat{h}} \))

  ∗ take \( x = \hat{g}, y = \hat{h}; \) \( \mathbb{X} := \{x : \|x\|_2 = 1\} \) (denoted by \( \text{IM}_{\hat{g}} \))

- **Kernel representation** (denoted by KER)
  Structured low-rank approximation \( \equiv \) (OUTER) with

\[
\begin{aligned}
f(x) &= \min_{y \in \mathbb{R}} \|y\|_2^2 \quad \text{subject to} \quad G(x) y = s(x), \quad \text{(LN)}
\end{aligned}
\]

where \( G(x), s(x) \) depend on the structure \( \gamma_2 \) and the polynomials \( p^k \). In this case, (LN) is a least norm problem.

### Optimization methods

Under some conditions on the weights,

\[
f(x) = \begin{cases} C - s(x)^{\top} \Gamma^{-1}(x)s(x), & \text{for } \text{IM}_{\hat{h}} \text{ and } \text{IM}_{\hat{g}} \text{.}
\end{cases}
\]

where \( s \) is a linear function and \( \Gamma \) is symmetric positive definite.

\( f \) is minimized with the SLRA package [3,4] for structured low-rank approximation (the Levenberg-Marquardt method is used).

SLRA package exploits structure and bandedness of \( \Gamma \) to evaluate the cost function, its first- and second-order derivatives [2].

### Features of the parameterizations

Denote \( n = \sum_{k=1}^{N} n_k, \ell = n - N d. \)

<table>
<thead>
<tr>
<th>Representation</th>
<th>IM</th>
<th>KER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complexity per iteration</td>
<td>( O(n^2) )</td>
<td>( O(n^2) )</td>
</tr>
<tr>
<td>Handles fixed values</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Handles missing values</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

† can handle, but with complexity \( O(n^3) \) per iteration.

Conclusions:

- Use \( \text{IM}_{\hat{h}} \) if \( d \ll n \) (small degree of the common divisor).
- Use \( \text{IM}_{\hat{g}} \) if \( (n_k - d) \ll n \) (small degree of the cofactors).

For more details, examples and comparisons — see [1].

The implementation is embedded in [1] using the literate program style. The methods are implemented in MATLAB and are based on the SLRA package [3,4].

### References


